

<https://helda.helsinki.fi>

A machine-learned analysis of human gene polymorphisms modulating persisting pain points to major roles of neuroimmune processes

Kringel, D.

2018-11

Kringel , D , Lippmann , C , Parnham , M J , Kalso , E , Ultsch , A & Lötsch , J 2018 , ' A machine-learned analysis of human gene polymorphisms modulating persisting pain points to major roles of neuroimmune processes ' , European Journal of Pain , vol. 22 , no. 10 , pp. 1735-1756 . <https://doi.org/10.1002/ejp.1270>

<http://hdl.handle.net/10138/256029>

<https://doi.org/10.1002/ejp.1270>

publishedVersion

Downloaded from Helda, University of Helsinki institutional repository.

This is an electronic reprint of the original article.

This reprint may differ from the original in pagination and typographic detail.

Please cite the original version.

ORIGINAL ARTICLE

A machine-learned analysis of human gene polymorphisms modulating persisting pain points to major roles of neuroimmune processes

D. Kringsel¹, C. Lippmann², M.J. Parnham², E. Kalso^{3,4}, A. Ultsch⁵, J. Lötsch^{1,2}

¹ Institute of Clinical Pharmacology, Goethe - University, Frankfurt am Main, Germany

² Fraunhofer Institute for Molecular Biology and Applied Ecology IME, Branch for Translational Medicine and Pharmacology TMP, Frankfurt

³ Institute of Clinical Medicine, University of Helsinki, Pain Clinic, Helsinki University Central Hospital, Helsinki, Finland

⁴ Institute of Biomedicine, Pharmacology, University of Helsinki, Helsinki, Finland

⁵ DataBionics Research Group, University of Marburg, Germany

Correspondence

Jörn Lötsch

E-mail: j.loetsch@em.uni-frankfurt.de

Funding sources

This work was funded by the European Union Seventh Framework Programme (FP7/2007 - 2013) under grant agreement no. 602919 (EK, JL, GLORIA) and in addition by the Landesoffensive zur Entwicklung wissenschaftlich-ökonomischer Exzellenz (LOEWE) Zentrum for Translational Medicine and Pharmacology project 'Process pharmacology: A data science based approach to drug repurposing' (JL). The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Conflicts of interest

The authors have declared that no further conflicts of interest exist.

Accepted for publication

13 June 2018

doi:10.1002/ejp.1270

Abstract

Background: Human genetic research has implicated functional variants of more than one hundred genes in the modulation of persisting pain. Artificial intelligence and machine-learning techniques may combine this knowledge with results of genetic research gathered in any context, which permits the identification of the key biological processes involved in chronic sensitization to pain.

Methods: Based on published evidence, a set of 110 genes carrying variants reported to be associated with modulation of the clinical phenotype of persisting pain in eight different clinical settings was submitted to unsupervised machine-learning aimed at functional clustering. Subsequently, a mathematically supported subset of genes, comprising those most consistently involved in persisting pain, was analysed by means of computational functional genomics in the Gene Ontology knowledgebase.

Results: Clustering of genes with evidence for a modulation of persisting pain elucidated a functionally heterogeneous set. The situation cleared when the focus was narrowed to a genetic modulation consistently observed throughout several clinical settings. On this basis, two groups of biological processes, the immune system and nitric oxide signalling, emerged as major players in sensitization to persisting pain, which is biologically highly plausible and in agreement with other lines of pain research.

Conclusions: The present computational functional genomics-based approach provided a computational systems-biology perspective on chronic sensitization to pain. Human genetic control of persisting pain points to the immune system as a source of potential future targets for drugs directed against persisting pain. Contemporary machine-learned methods provide innovative approaches to knowledge discovery from previous evidence.

Significance: We show that knowledge discovery in genetic databases and contemporary machine-learned techniques can identify relevant biological processes involved in Persistent pain.

1. Introduction

Persisting pain has a high prevalence (Elliott et al., 1999; Breivik et al., 2006; van Hecke et al., 2013) affecting a significant proportion of the world's population (Breivik et al., 2006). Its pathophysiology is incompletely understood, which is reflected in the limited success of available treatment options (Moore et al., 2009, 2012; Derry et al., 2012) and has stimulated intense research on this topic (Kringel and Lötsch, 2015). In this context, the study of human gene polymorphisms that modulate the persisting pain phenotype is an accepted research approach which has been pursued for more than 50 years (Godinova, 1965). A genetic background to persisting pain is clearly reflected by a protective effect against persisting pain exerted, for example, by a haplotype of the guanosine-5'-triphosphate (GTP) cyclohydrolase 1 gene (*GCH1*), originally reported to be composed of 15 genetic variants (Tegeder et al., 2006, 2008), or by a reduction in the perceived intensity of pain exerted, for example, by a deletion/insertion variant in the serotonin transporter gene-linked polymorphic region (5-HTTLPR; gene: *SLC6A4*) reportedly reducing the perception of heat pain (Lindstedt et al., 2011; Hooten et al., 2013; Kunz et al., 2016). On the other hand, increased risk for persisting pain is conferred, for example, by the rs12584920 variant of the 5-hydroxytryptamine receptor 2A gene (*HTR2A*) (Nicholl et al., 2011) or the rs734784 polymorphism in the potassium voltage-gated channel modifier, subfamily S member 1, gene (*KCNS1*) (Costigan et al., 2010).

Human genetic research during the last decade has identified many common variants of more than a hundred different genes spread across the genome that modulate the phenotype of persisting pain in several different clinical settings (Table 1). Thanks to concomitant developments in computer science, including progress in artificial intelligence, machine-learning and knowledge discovery in databases (Ashburner et al., 2000), the analysis of fundamental, complex biological functions has become increasingly possible. This allows persisting pain to be approached at a functional genomics level by combining the information on genetic modulation acquired in clinical studies with current knowledge of the function of human genes. This active research topic has already led to the identification of candidate genes for further clinical genetic pain research (Lötsch et al., 2013) and highlighted key pathophysiological processes of pain which may be targeted for future pharmacological treatment options (Ultsch et al., 2016).

In the present analysis, information on genes, for which empirical evidence indicates the existence of variants that modulate the clinical phenotype of persisting pain, was analysed at a functional genomics level. In this way, the biochemical, cellular and/or physiological properties of each and every gene product can be investigated to gain an understanding of the relationship between the genome and the phenotype on a global genomewide scale (Gibson and Muse, 2009). Applying machine-learned techniques (Fig. 1), the genes currently known to have relevance to human persisting pain were analysed for functional patterns that may provide insight into its pathophysiology based on current research activities, applying a data-driven approach without using prior hypotheses about the most important biological functions characterizing persisting pain. By applying methods of precisely calculated item selection (Ultsch and Lötsch, 2015), the present analysis aimed to identify a subset of genes that were most consistently reported to be involved in the modulation of persisting pain with a subsequent analysis of the main biological functions exerted by the products of these genes.

2. Methods

2.1 Search strategy

A set of genes relevant to persisting pain, based on published associations of their variants with phenotypic differences in persisting pain patients, was obtained from (1) a PubMed database (accessed in August 2016), by searches for the string '(chronic OR persisting OR neuropathic OR back OR inflammatory OR musculoskeletal OR visceral OR widespread OR idiopathic OR fibromyalgia) AND pain AND (polymorphism OR variant) NOT review' and (2) publications starting from the year 2000, which is close to the first sequencing of the human genome (Lander et al., 2001; Venter et al., 2001) marking the beginning of a new area of genetic research, and (3) published overviews on pain genetics (e.g. Edwards, 2006; Tegeder and Lötsch, 2009; Mogil, 2012; Zorina-Lichtenwalter et al., 2016). To avoid redundancies, reports of positive associations of any gene variant were included only once per clinical setting in the present analysis. This implies that not every variant found to be functionally associated with a persisting pain phenotype was taken into account. The resulting information for each gene, thus, comprised (1) a positive report of a gene

Table 1 List of 110 genes with empirically supported relevance to persisting pain, based on published evidence that their genetic variants are associated with phenotypic differences in persisting pain patients in several clinical settings.

Gene	NCBI	Type of pain	References	Gene	NCBI	Type of pain	References	Gene	NCBI	Type of pain	References
ACAN	176	Musculoskeletal	Kirk et al. (2003)	GCH1	2643	Widespread	Kim et al. (2010a)	MTHFD1	4522	Musculoskeletal	Aneiros-Guerrero et al. (2011)
ACE	1636	Musculoskeletal	Rommel et al. (2006)	GDF5	8200	Musculoskeletal	Valdes et al. (2011c)	MTRR	4552	Musculoskeletal	Aneiros-Guerrero et al. (2011)
ADRA1A	148	Neuropathic	Herlyn et al. (2010)	GNB3	2784	Visceral	Oshima et al. (2010)	MYT1L	23,040	Widespread	Docampo et al. (2014)
ADRA1D	146	Visceral	Sugaya et al. (2002)	GRK5	2869	Miscellaneous	Smith et al. (2011)	NCR3	259,197	Neuropathic	Sato et al. (2002)
ADRA2A	150	Idiopathic	Kim et al. (2004)	GSTM1	2944	Visceral	Wu et al. (2000)	NPY	4852	Back pain	Herlyn et al. (2010)
ADRA2C	152	Idiopathic	Kim et al. (2004)	GSTM1	2944	Musculoskeletal	Aneiros-Guerrero et al. (2011)	NR3C1	2908	Widespread	Holliday et al. (2010)
ADRB2	154	Widespread	Hocking et al. (2010)	GSTP1	2950	Visceral	Woo et al. (2010)	NRXN3	9369	Widespread	Docampo et al. (2014)
ADRB2	154	Musculoskeletal	Diatchenko et al. (2006)	GSTT1	2952	Visceral	Woo et al. (2010)	NTRK1	4914	Miscellaneous	Shatzky et al. (2000)
ADRB2	154	Neuropathic	Herlyn et al. (2010)	HFE	3077	Musculoskeletal	Alizadeh et al. (2007)	OPRM1	4988	Miscellaneous	Cheng et al. (2010)
ADRB3	155	Visceral	Sugaya et al. (2002)	HLA-A	3105	Neuropathic	Sato et al. (2002)	P2RX7	5027	Musculoskeletal	Sorge et al. (2012)
ANP32A	8125	Musculoskeletal	Valdes et al. (2009)	HLA-B	3106	Neuropathic	de Rooij et al. (2009)	P2RX7	5027	Neuropathic	Sorge et al. (2012)
APOE	348	Widespread	Reeser et al. (2011)	HLA-B	3106	Inflammatory	Gullo et al. (1982)	PCSK6	5046	Musculoskeletal	Malfait et al. (2012)
AR	367	Visceral	Shaik et al. (2009)	HLA-C	3107	Neuropathic	Ozawa et al. (1999)	PGK1	5230	Visceral	Riley and Krieger, 2002)
ASPN	54,829	Musculoskeletal	Nakamura et al. (2007)	HLA-DQA1	3117	Neuropathic	de Rooij et al. (2009)	PGR	5241	Visceral	De Carvalho et al. (2007)
CACNA2D3	55,799	Back pain	Neely et al. (2010)	HLA-DQB1	3119	Neuropathic	de Rooij et al. (2009)	POMC	5443	Widespread	Holliday et al. (2010)
CACNG2	10,369	Neuropathic	Nissenbaum et al. (2010)	HLA-DRB1	3123	Neuropathic	Sato et al. (2002)	PRSS1	5644	Inflammatory	Midha et al. (2010)
CALCA	796	Neuropathic	Herlyn et al. (2010)	HTR2A	3356	Visceral	Pata et al. (2004)	PTGS1	5742	Visceral	Arisawa et al. (2008)
CAMK4	814	Miscellaneous	Smith et al. (2011)	HTR3E	285,242	Visceral	Kilpatrick et al. (2011)	SCN5A	6331	Visceral	Saito et al., 2009)
CASP9	842	Back pain	Guo et al. (2011)	IFNG	3458	Inflammatory	Noponen-Hietala et al. (2005)	SCN5A	6331	Idiopathic	Reimann et al. (2010)
CCT5	22,948	Widespread	Peters et al. (2013)	IFNG	3458	Inflammatory	Oen et al. (2005)	SCN9A	6335	Back pain	Reimann et al. (2010)
CFTR	1080	Inflammatory	Midha et al. (2010)	IFRD1	3475	Miscellaneous	Smith et al. (2011)	SCN9A	6335	Inflammatory	Reimann et al. (2010)

Table 1 (Continued)

Gene	NCBI	Type of pain	References	Gene	NCBI	Type of pain	References	Gene	NCBI	Type of pain	References
<i>CRHBP</i>	1393	Widespread	Holliday et al. (2010)	<i>IL-10</i>	3586	Back pain	Shoskes et al. (2002)	<i>SCN9A</i>	6335	Miscellaneous	Reimann et al. (2010)
<i>CILP</i>	8483	Back pain	Seki et al. (2005)	<i>IL-10</i>	3586	Inflammatory	Noponen-Hietala et al. (2005)	<i>SCN9A</i>	6335	Musculoskeletal	Valdes et al. (2011a)
<i>CNR1</i>	1268	Visceral	Park et al. (2011)	<i>IL-10</i>	3586	Musculoskeletal	Oen et al. (2005)	<i>SCN10A</i>	6336	Neuropathic	Faber et al. (2012)
<i>COL1A1</i>	1277	Back pain	Tilkeridis et al. (2005)	<i>IL-16</i>	3603	Visceral	Gan et al. (2010)	<i>SCN11A</i>	11,280	Widespread	Leipold et al., (2013)
<i>COL6A4P1</i>	344,875	Musculoskeletal	Miyamoto et al. (2008)	<i>IL-1A</i>	3552	Back pain	Solovieva et al. (2004)	<i>SERPINA1</i>	5265	Widespread	Blanco et al. (2006)
<i>COL9A2</i>	1298	Back pain	Ala-Kokko, 2002)	<i>IL-1B</i>	3553	Back pain	Zhang et al. (2002)	<i>SERPINA6</i>	866	Widespread	Holliday et al. (2010)
<i>COL9A3</i>	1299	Back pain	Kales et al. (2004)	<i>IL-1B</i>	3553	Miscellaneous	Jeremias et al. (2000)	<i>SHMT1</i>	6470	Idiopathic	Aneiros-Guerrero et al. (2011)
<i>COMT</i>	1312	Musculoskeletal	van Meurs et al. (2009)	<i>IL-1R2</i>	7850	Neuropathic	Stephens et al. (2014)	<i>SLC6A4</i>	6532	Idiopathic	Herken et al. (2001)
<i>COMT</i>	1312	Widespread	Cohen et al. (2009)	<i>IL-1RN</i>	3557	Back pain	Kim et al. (2010b)	<i>SMAD3</i>	4088	Musculoskeletal	Valdes et al. (2010)
<i>COMT</i>	1312	Back pain	Dai et al. (2010)	<i>IL-1RN</i>	3557	Musculoskeletal	Attur et al. (2010)	<i>SOD2</i>	6648	Idiopathic	Arisan et al. (2006)
<i>COMT</i>	1312	Visceral	Karling et al. (2011)	<i>IL-1RN</i>	3557	Idiopathic	Witkin et al. (2002)	<i>SPINK1</i>	6690	Inflammatory	Midha et al. (2010)
<i>COMT</i>	1312	Idiopathic	Tahara et al. (2008)	<i>IL-2</i>	3558	Inflammatory	Noponen-Hietala et al. (2005)	<i>TAAR1</i>	134,864	Widespread	Smith et al. (2012)
<i>CRH</i>	1392	Widespread	Holliday et al. (2010)	<i>IL-4</i>	3565	Visceral	Sugaya et al. (2002)	<i>TAC1</i>	6863	Back pain	Herlyn et al. (2010)
<i>CRHBP</i>	1393	Musculoskeletal	Linnstaedt et al. (2016)	<i>IL-4</i>	3565	Inflammatory	Noponen-Hietala et al. (2005)	<i>TACR1</i>	6869	Visceral	Renner et al. (2009)
<i>CRHR1</i>	1394	Widespread	Holliday et al. (2010)	<i>IL-4R</i>	3566	Visceral	Sugaya et al. (2002)	<i>TGFB1</i>	7040	Back pain	Herlyn et al. (2010)
<i>CYP2D6</i>	1565	Visceral	Wu et al. (2000)	<i>IL-4R</i>	3566	Inflammatory	Noponen-Hietala et al. (2005)	<i>TGFB1</i>	7040	Inflammatory	Shoskes et al. (2002)
<i>DIO2</i>	1734	Musculoskeletal	Meulenbelt et al. (2008)	<i>IL-6</i>	3569	Back pain	Herlyn et al. (2010)	<i>TGFB1</i>	7040	Musculoskeletal	Oen et al. (2005)
<i>DRD4</i>	1815	Musculoskeletal	Aneiros-Guerrero et al. (2011)	<i>IL-6</i>	3569	Idiopathic	Shoskes et al. (2002)	<i>TNF</i>	7124	Back pain	Herlyn et al. (2010)
<i>DRD4</i>	1815	Widespread	Buskila et al. (2004)	<i>IL-6</i>	3569	Inflammatory	Noponen-Hietala et al. (2005)	<i>TNF</i>	7124	Idiopathic	Shoskes et al. (2002)
<i>LPAR1</i>	1902	Musculoskeletal	Mototani et al. (2008)	<i>IL-6</i>	3569	Neuropathic	Oen et al. (2005)	<i>TNF</i>	7124	Inflammatory	Noponen-Hietala et al. (2005)
<i>ESR1</i>	2099	Idiopathic	Ribeiro-Dasilva et al. (2009)	<i>KCNJ6</i>	3763	Back pain	Bruehl et al. (2013)	<i>TNF</i>	7124	Musculoskeletal	Oen et al. (2005)

Table 1 (Continued)

Gene	NCBI	Type of pain	References	Gene	NCBI	Type of pain	References	Gene	NCBI	Type of pain	References
ESR1	2099	Visceral	Govindan et al. (2009)	KCN51	3787	Neuropathic	Costigan et al. (2010)	TP53	7157	Visceral	Ribeiro Junior et al. (2009)
ESR1	2099	Musculoskeletal	Kang et al. (2007)	MAOA	4128	Musculoskeletal	Gursoy et al. (2008)	TPH2	121,278	Widespread	Nicholl et al. (2011)
FAM173B	134,145	Widespread	Peters et al. (2013)	MBL2	4153	Idiopathic	Babula et al. (2004)	TPH2	121,278	Musculoskeletal	Nicholl et al. (2011)
FKBP5	2289	Musculoskeletal	Bortsov et al. (2013)	MCT1R	4157	Idiopathic	Foster et al. (2004)	TRPA1	8989	Neuropathic	Binder et al. (2011)
GBP1	2633	Widespread	Smith et al. (2012)	MC2R	4158	Widespread	Holliday et al. (2010)	TRPM8	79,054	Neuropathic	Binder et al. (2011)
GC	2638	Visceral	Faserl et al. (2011)	MIF	4282	Visceral	Arisawa et al. (2007)	TRPV1	7442	Musculoskeletal	Valdes et al. (2011b)
GCH1	2643	Back pain	Doehring et al. (2009)	SOD2	6648	Visceral	Arisan et al. (2006)	TRPV1	7442	Neuropathic	Binder et al. (2011)

The gene, NCBI number, clinical setting of persisting pain are shown together with a key reference in which this association was reported. The studies are given grouped for the relevant gene, however, in arbitrary order of clinical settings or publication year.

modulation in persisting pain and (2) the clinical setting of this finding.

2.2 Data analysis

Data were analysed using the R software package (version 3.3.2 for Linux; <http://CRAN.R-project.org/>; R Development Core Team, 2008) on an Intel Xeon® computer running on Ubuntu Linux 16.04.1 64-bit. The analysis employed several methods of machine-learning that, as described previously (Lötsch and Ultsch, 2017a), may be referred to as a set of methods that can automatically detect patterns in data and then use the uncovered patterns to predict or classify future data, to observe structures such as subgroups in the data or to extract information from the data suitable to derive new knowledge (Murphy, 2012; Dhar, 2013). More detailed descriptions including definitions of key concepts have been provided elsewhere (Lötsch and Ultsch, 2017a).

The analysis aimed at describing the functional genomics of persisting pain based on the biological roles of the genes that reportedly carry variants modulating that phenotype. The biological roles were assessed as biological processes in which the genes are involved, defined as series of events or molecular functions with a defined beginning and end (Ashburner et al., 2000) and queried from the Gene Ontology (GO) knowledgebase that provides the acquired knowledge about the biological functions of gene products, described with a controlled vocabulary of GO terms (Ashburner et al., 2000).

Thus, the basis on which the present functional picture of persisting pain was created consisted of the biological processes in which the genes carrying modulatory variants were reported to be involved. The functional picture of persisting pain was sought pursuing two different analytical paths (Fig. 1). In a *first* approach, functional subgroups were sought in the set of human genes, variants of which have been associated with modulation of the clinical phenotype of persisting pain. This was addressed by applying a clustering algorithm on the matrix given by the genes versus their annotated biological processes; an approach that previously proved as suitable for gene function based classifications (Lötsch and Ultsch, 2016a). In a *second* approach, the hypothesis was pursued that the functional genomics of persisting pain will prevail across clinical settings irrespective of the disease that had originally triggered the process. Therefore, the most informative subsets of the genes were identified using a computed item categorization technique

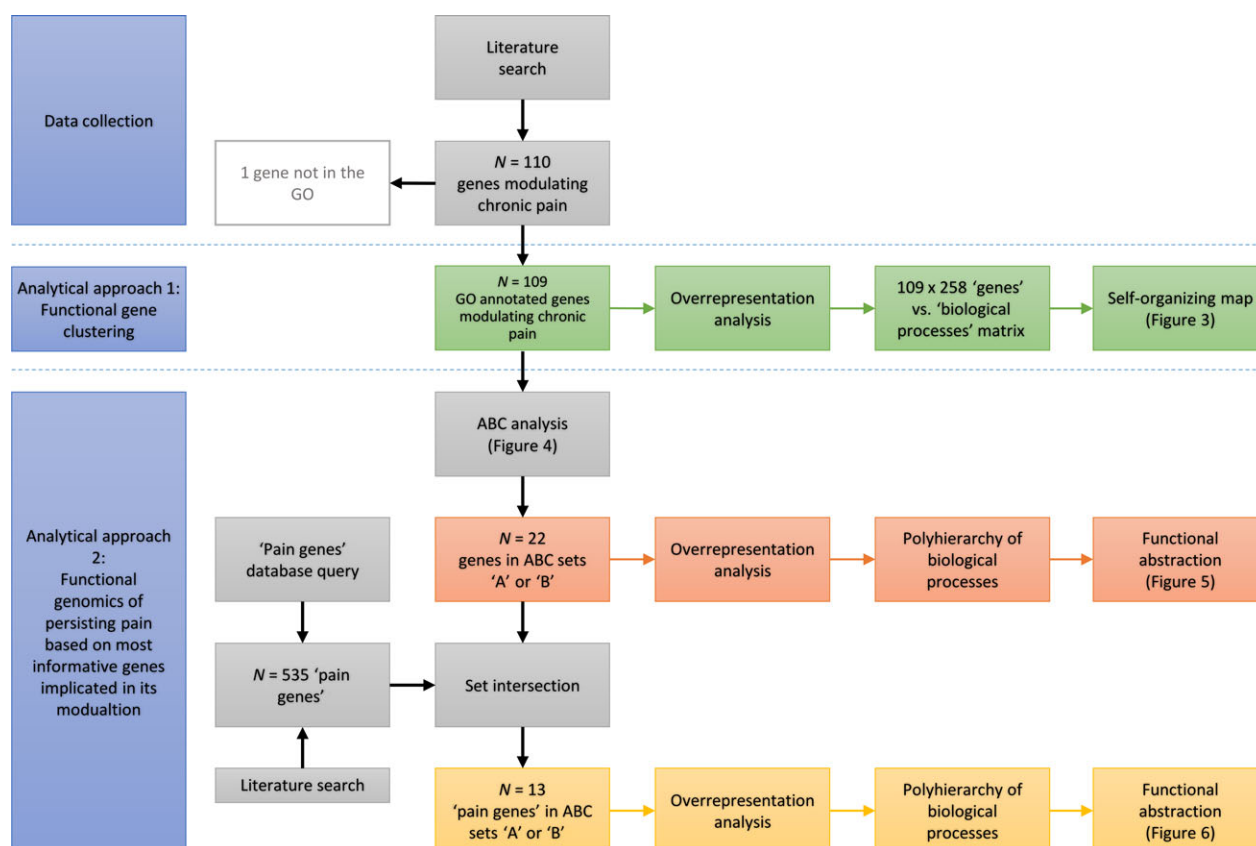


Figure 1 Flow chart of the data analysis. The figure provides an overview of the machine-learning and data science approach applied, specifying the data flow of the gene sets and the analyses applied to them. The figure follows the analytical flow (left column of blue rectangles, separated by horizontal dashed lines) that following data collection from the literature was implemented as two analytical approaches comprising (1) functional gene clustering and (2) a computational functional genomics analysis of the biological roles of the genes implicated in the modulation of persisting pain in various clinical settings. The coloured frames show the presented analysis, the grey frames the intermediate data flows.

(Ultsch and Lötsch, 2015) and subsequently, techniques of knowledge discovery were applied to identify the particular biological roles exerted by this set of genes as opposed to the biological functions exerted by a similarly sized random set of genes. This analysis was performed twice, once using the most informative genes resulting from above-mentioned analysis, and, to strengthen the evidence for pain reliance, again using the additional selection criterion that the genes should be listed among pain-relevant genes in the PainGenes database (Lacroix-Fralish et al., 2007). The analytical steps are described in detail in the following.

2.2.1 Functional clustering of genes

A first analytical approach to the functional genomics of persisting pain reflected in the set of genes that carry variants reported to modulate the phenotype aimed at finding functional clusters of genes.

2.2.2 Generation of a gene versus functional feature matrix

As a prerequisite for functional clustering of genes, a gene versus function matrix was created. Therefore, the biological functions in which the genes, or their products, are involved were queried from the GO knowledge base (<http://www.geneontology.org/>). The GO knowledge base is searchable for three major categories, consisting of biological process, cellular component and molecular function. As the most suitable representation of processes that are involved in the chronification of central sensitization to pain, the GO category biological process was selected as previously (Lötsch et al., 2013; Lötsch and Ultsch, 2016b; Ultsch et al., 2016). According to the GO knowledgebase, this category contains one or more ordered collections of molecular functions involving chemical or physical transformations such as cell growth and maintenance or signal transduction (Ashburner et al., 2000).

However, not all processes known to be influenced by the genes were sought, but only those that were annotated to the present set of genes more often than expected for any similarly sized random set of genes. Therefore, to capture biological processes that are particularly relevant to the present gene set, the set of genes was submitted to overrepresentation analysis (ORA; Backes et al., 2007). As intended, this compared the occurrence (as defined by its annotation term) of the particular set of genes covered by a GO term with the number of genes expected to be defined by this term. The significance of the association of a GO term with the expected list of genes was determined by means of a Fisher's exact test (Fisher, 1922). A p -value threshold, t_p , of 1×10^{-6} was applied and subsequently, the obtained results were additionally corrected for multiple testing according to Bonferroni (Hochberg, 1988). The result was the desired 'gene versus biological process' matrix (Fig. 2) that, rescaled as [0,1] indicating the absence or presence, respectively, of the involvement of a gene in a particular biological process, provided a filtered representation of the particularly important processes in which the analysed genes were involved while disregarding processes that would have been found by chance in any similarly sized gene set. This ORA-based filtering of gene functions was previously found to facilitate the functional analysis of gene sets including a context of pain and analgesia (Lötsch and Ultsch, 2016a; Lötsch et al., 2017).

2.2.3 Machine-learned cluster detection

Following the creation of the gene versus functional feature matrix, expressed as 'gene versus biological process' matrix, the feature space required for functional gene clustering was established as $D = \{(x_i) | x_i \in N^d, i = 1, \dots, n\}$ comprising the d biological process to which the n genes in the analysed set were annotated. This feature space was searched for a cluster structure. Among several methods available for clustering, a method of unsupervised machine-learning shown recently to provide a viable unbiased clustering of high-dimensional biomedical data, outperforming classical clustering algorithms (Ultsch and Lötsch, 2017), was chosen. Specifically, the data space was projected from the high-dimensional feature space D onto a two-dimensional self-organizing map (SOM) of the Kohonen type (Kohonen, 1982). This map was composed of a toroid grid (Ultsch, 2003), i.e. a projection plane where opposite edges are connected. The grid had a size of 25×35 artificial neurons chosen according to the proposals of SOM

size determination described previously (Ultsch and Lötsch, 2017). Each of the artificial neurons held a position vector, which carried the information about the biological processes associated with each gene, and a further parameter, which carried 'weights' of the same dimensions as the input dimensions. The weights were initially randomly drawn from the range of the data variables and subsequently adapted to the data during the learning phase of 25 epochs. The Euclidean distance was used for process (dis-)similarity; very general processes, such as 'biological process' that is the root term of the polyhierarchy carry the same value for all genes and therefore do not influence this distance. Following training of the neural network, an emergent SOM (ESOM; Lötsch and Ultsch, 2014; Ultsch and Sieman, 1990) was obtained that represented the genes as the localizations of their 'best matching units' (BMU), which are neurons that carried the vector most similar to a gene's data vector.

Following the projection of the data on the grid of neurons, an extension of the Kohonen map was applied to obtain clusters of genes. Specifically, the distance structure in the high-dimensional space was visualized using the so-called U-matrix (Ultsch and Sieman, 1990; Lötsch and Ultsch, 2014). The clusters became visible using a geographical map analogy where 'mountain ranges' represent large distances in the feature space that can be used to visually separate data clusters, whereas low 'valleys' represent sets of genes that are related to similar biological processes and therefore belong to the same cluster. The 'map' was further enhanced by calculating a so-called P-matrix (Ultsch, 2003), which displays the probability of a data point as $p(x) = |\{ \text{data points } x_i | d(x_i, x) \leq r \}|$ estimated as the number of data points in a sphere with radius r around x at each grid point on the ESOM's output grid. The calculations were performed using our R library 'Umatrix' (<https://cran.r-project.org/package=Umatrix>; Lötsch et al., 2018a).

2.2.4 Functional genomics analysis of most informative genes reported to modulate persisting pain

A second analytical approach at the functional genomics of persisting pain reflected in the set of genes that carry variants reported to modulate the phenotype aimed at identifying the biological roles of those genes that had been implicated most consistently in this context. Specifically, the present analysis aimed at the discovery of new knowledge about pain, rather than about the underlying disease, from

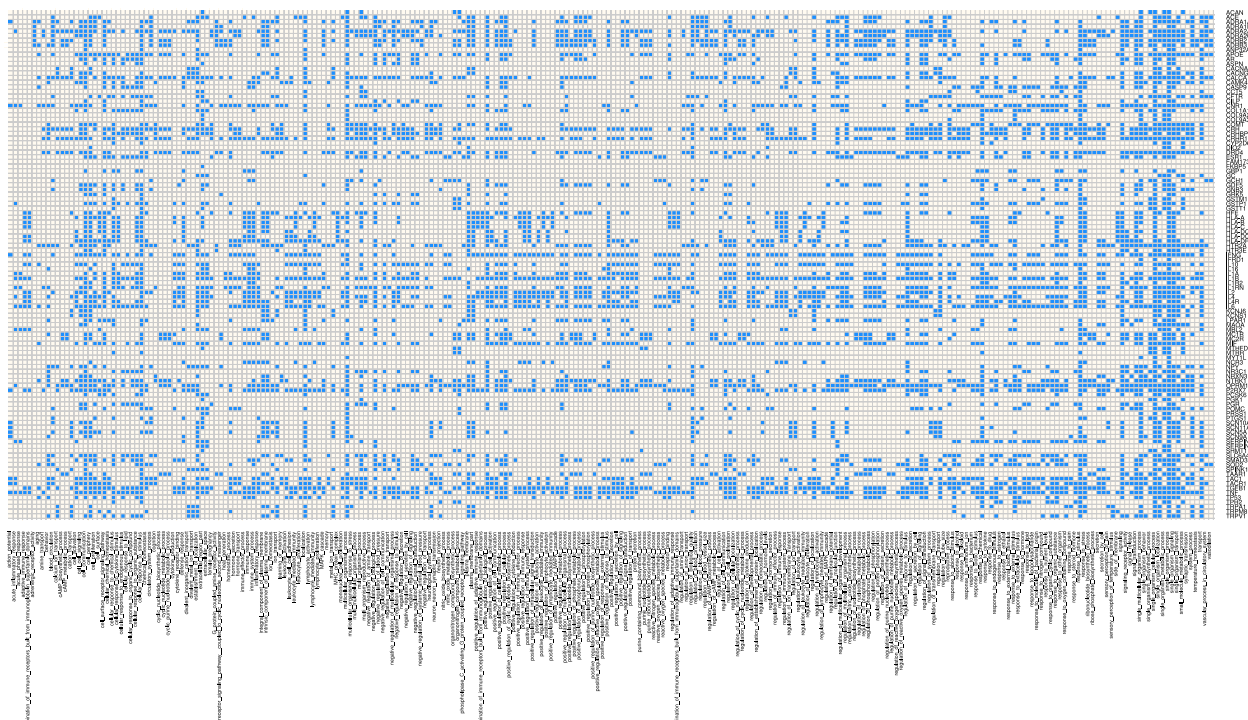


Figure 2 Matrix plot of the association of identified genes (rows; Table 1) with biological processes (columns), according to the annotations of the Gene Ontology (GO) knowledgebase (Ashburner et al., 2000), filtered for statistical significance in the present context by means of an overrepresentation analysis with a p -value threshold of 1×10^{-6} and correction for multiple testing according to Bonferroni. The matrix displays a yes/no scaling [1,0], colour-coded as blue or white, respectively. The figure has been created using the R software package (version 3.3.2 for Linux; <http://CRAN.R-project.org/>; R Development Core Team, 2008). Specifically, this plot was created using the 'heatmap.2' functions of the R package 'gplots' (Warnes G. R.; <https://cran.r-project.org/package=gplots>) with the build-in clustering of the plotting routine disabled (R switches 'Colv = FALSE, Rowv = FALSE'). For descriptions of the GO terms (abscissa), see the AmiGO search tool at <http://amigo.geneontology.org/> (Carbon et al., 2009).

reports about a genetic modulation of a clinical symptom involving pain. This implies a distinction between the modulation of pain and the modulation of the disease-causing pain that has usually not been made in the included reports. For example, in rheumatic diseases a genetic variant could modulate the progress and severity of inflammatory processes and via this it could finally modulate pain, or a genetic variant could directly modulate the individual perception of pain and therefore, similar nociceptive stimuli produced by the inflammatory processes could cause different degrees of pain. However, in the interest of the present analysis, we assessed only the direct modulations of pain, which were expected to be provided most likely by those genes that had been implicated in modulation of persisting pain in more than one clinical setting.

2.2.5 Identification of genes most consistently reported to modulate persisting pain

To identify the most informative subset of genes reportedly modulating clinical persisting pain, a

cut-off criterion had to be defined for the number of different clinical settings required for inclusion in further analyses. To avoid arbitrary criteria, the cut-off was obtained by applying an item categorization technique used to separate 'the important few' from the 'trivial many' (Juran, 1975). As a most suitable technique, because providing a mathematically based cut-off, computed ABC analysis was chosen for the present purpose, supported by previous demonstrations that it is suitable for item selection tasks in biomedical research (Ultsch and Lötsch, 2015; Lötsch and Ultsch, 2017b; Lötsch et al., 2018b).

ABC analysis requires a set of positive numbers, which was given by the column sums of the 'clinical settings versus genes' matrix (Fig. 4 top). The vector of the sums of clinical settings with positive reports of the modulatory involvement a gene's variant was submitted to computed ABC analysis (Ultsch and Lötsch, 2015), which aims to divide a set of positive data – here the set of genes scored according to their involvement in clinical settings

of persisting pain - into three disjointed subsets called 'A', 'B' and 'C'. Subset 'A' comprises 'the important few', subset 'C' comprises clearly non-profitable values, i.e. 'the trivial many' (Juran, 1975), whereas subset 'B' includes items that provide still a balance between effort and gain. Therefore, the limit separating subset 'C', i.e. the genes for which a modulation of persisting clinical pain provides the least relevant information, was chosen as the limit for the inclusion of genes in further analyses. These calculations were made using our R package 'ABCanalysis' (<https://cran.r-project.org/package=ABCanalysis>; Ultsch and Lötsch, 2015).

2.2.6 Functional genomics analysis of genes most consistently reported to modulate persisting pain

Following identification of the most informative genes as members of ABC sets 'A' or 'B', the functional genomics picture of persisting pain arising from these genes was analysed. This was obtained by applying ORA as described above, again using a p -value threshold of $t_p = 1 \times 10^{-6}$ with Bonferroni α -correction for multiple testing. The focus of this analysis was, however, the hierarchical representation of the complete knowledge on the biological roles of genes that carry polymorphisms observed to modulate persisting pain phenotypes. This was provided in a directed acyclic graph (DAG; Thulasiraman and Swamy, 1992). In this graph, the top-down, branching polyhierarchy of GO terms starts with the most broadly defined terms and progresses towards the branches, representing GO terms with the narrowest definition (details). These calculations were made using our R package 'dbtORA' (<https://github.com/IME-TMP-FFM/dbtORA>; Lippmann et al., 2018), which has been designed for knowledge discovery in the GO.

As the complete DAG usually contains many GO terms (e.g. 64 GO terms in the present ORA), the information was transformed into a more intelligible form using the method of 'functional abstraction' (Ultsch and Lötsch, 2014). This aims to reduce the numbers of GO terms using a heuristic search algorithm that identifies so-called functional areas (Ultsch and Lötsch, 2014), which are GO terms that qualify by their informational importance as headlines representing specific aspects (taxonomies) of the complete DAG with maximal coverage, precision, informational value and conciseness (Ultsch and Lötsch, 2014).

To narrow the focus even more to pain-relevant genes, the ORA was repeated using the set

intersection of the most informative genes identified as described above by means of ABC analysis, with the genes listed among known pain-relevant genes in the PainGenes database (<http://www.jbldesign.com/jmogil/enter.html>; Lacroix-Fralish et al., 2007). These mainly include genes found, in at least three independent studies in transgenic mice, to contribute to the modulation of pain and identified using PubMed searches, with the addition of further genes (Lötsch et al., 2013) comprising those causally implicated in human hereditary diseases associated with extreme pain phenotypes (summarized in, e.g. Lötsch et al., 2017), and genes coding for the targets of approved analgesic drugs or of novel analgesics currently in clinical phases of development (Lötsch and Geisslinger, 2011). This provided a set of $n = 535$ 'pain genes' (Fig. 1 bottom).

3. Results

As a result of a literature search, a total of 110 unique genes were identified in eight different clinical settings of chronic central sensitization to pain, including back pain, inflammatory pain, musculoskeletal pain, neuropathic pain, visceral pain, widespread pain, idiopathic pain and miscellaneous pain, for which functional associations of genetic variants with differences in the phenotype of persisting pain had been reported (Table 1). Some of the included studies used a genomewide approach; however, many were candidate gene studies.

3.1 Functional clustering of genes carrying variants reportedly modulating persisting clinical pain

In a first analytical approach, functional subgroups were sought in the set of human genes, variants of which have been associated with modulation of the clinical phenotype of persisting pain. A filtered representation of the particularly important processes in which the analysed genes were involved while disregarding processes that would have been found by chance in any similarly sized gene set, was obtained by means of overrepresentation analysis (ORA) of the biological processes to which the genes were annotated in the GO knowledgebase. This identified $d = 258$ biological processes (GO terms), given the p -value threshold of 1×10^{-6} and the α -correction according to Bonferroni. One gene was neglected in this analysis, *COL6A4P1*, the collagen type VI alpha 4 pseudogene 1, because it was not referenced in the GO.

Subsequently, the 109×258 sized 'gene to biological process matrix' (Fig. 2) thus obtained was analysed for functional subgroups of genes. Using unsupervised machine-learning implemented as self-organizing artificial neuronal network of the Kohonen type enhanced by the U-matrix (i.e. an emergent self-organizing map, ESOM; Ultsch and Sieman, 1990; Lötsch and Ultsch, 2014), the high-dimensional data space was projected onto a two-dimensional toroid grid. On this grid, the U-matrix was visualized by applying a geographical landscape analogy (Fig. 3) providing a visual structure that could be employed for clustering of genes. The results indicated a large cluster comprising more than half of the genes ($n = 58$). However, this cluster was still composed of functionally very different genes, pointing towards a large heterogeneity of the genes chosen as candidate modulators of persisting pain in the different clinical studies, or was found with genomewide association studies (GWAS) without a focus on a particular gene. In addition, six smaller clusters were suggested, but their distinct separation was occasionally incomplete or, according to the U-matrix landscape analogy, they were not clearly presented as 'valleys' but consisted merely of 'mountain' zones separated by slightly higher ridges (Fig. 3). In ESOM/U-matrix based clustering, this indicates rather large intracluster distances.

As the present method has been shown to be well able to identify existent cluster structures while being unlikely to show false clusters (Ultsch, 2005; Ultsch and Lötsch, 2017), the main result of this analysis was that there is considerable heterogeneity among the genes reported to be involved in persisting pain without a clear functional focus on common general processes underlying this trait.

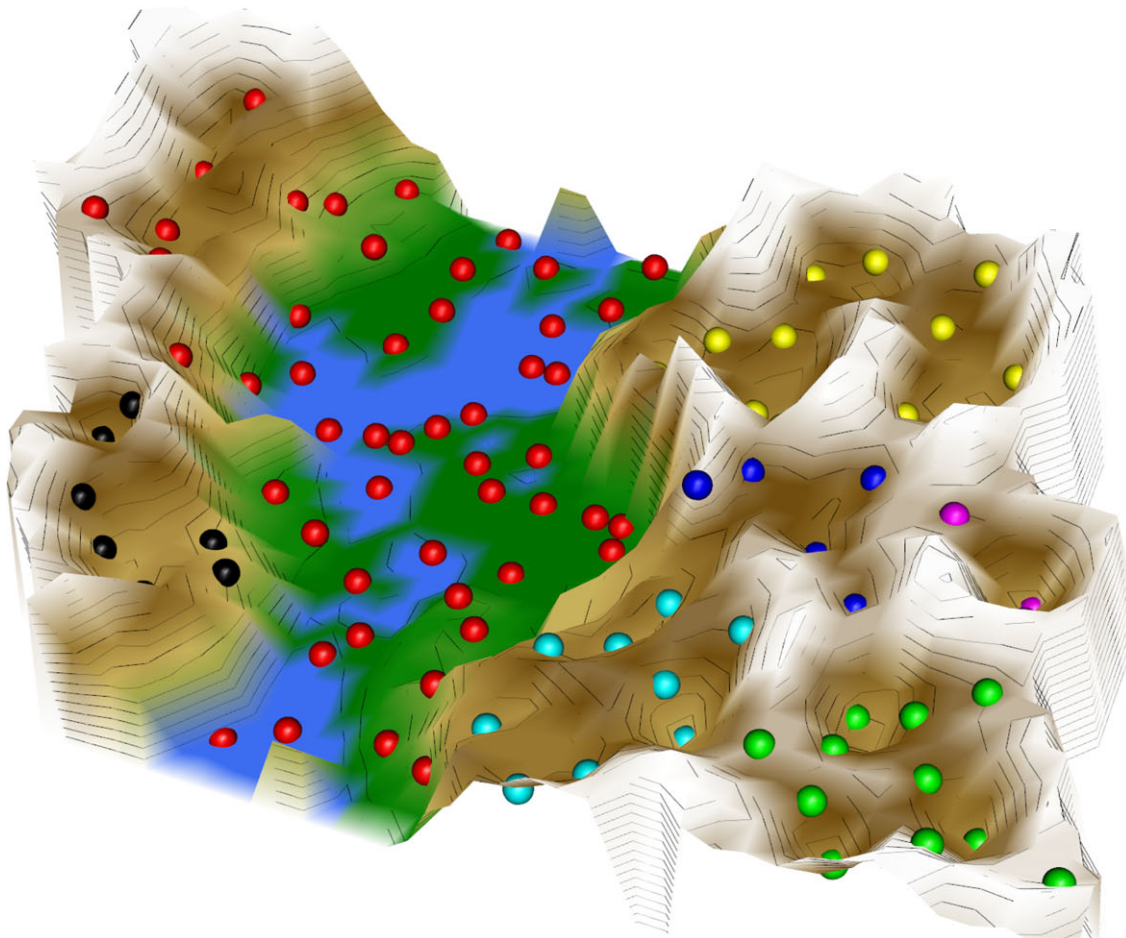
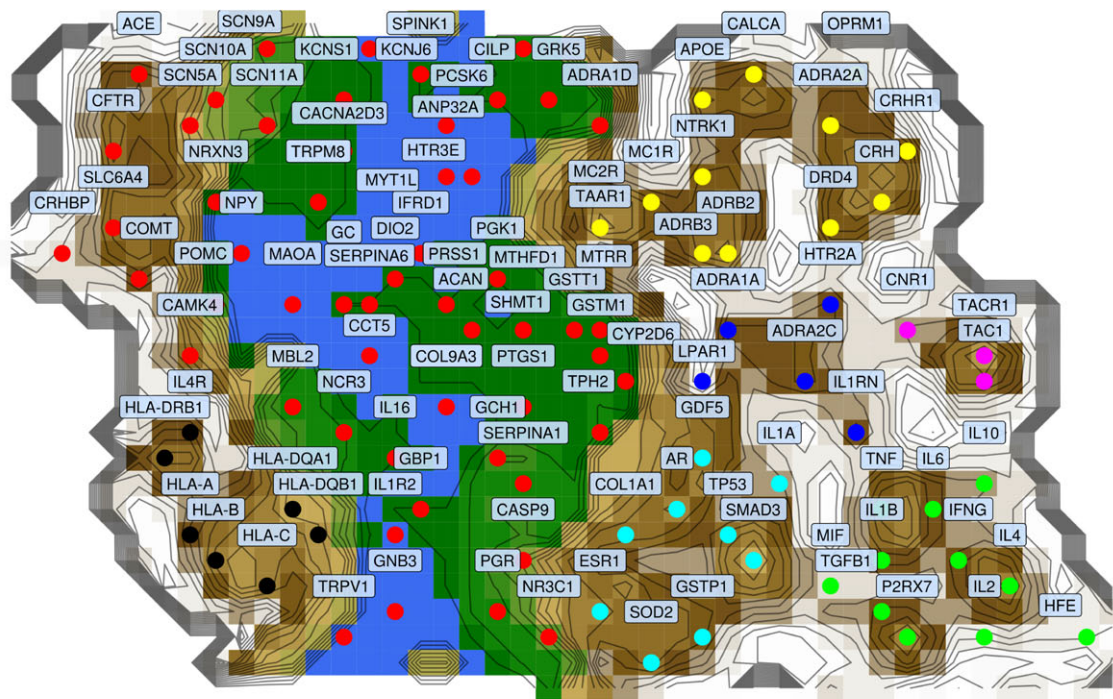
3.2 Functional genomics analysis of most informative genes reported to modulate persisting pain

In a second analytical approach, the hypothesis was pursued that the functional genomics of persisting pain will prevail across clinical settings irrespective of the disease that had originally triggered the process. However, the set intersection of the genes associated with any of the eight different clinical settings was empty. Nevertheless, the hypothesis that modulatory effects on chronic central sensitization to pain, rather than on the underlying disease-causing pain, are exerted by the same genes across several clinical settings could be pursued in several genes have been shown to be involved in more than one clinical setting of chronic central sensitization.

To identify how many clinical settings qualify as a cut-off, an ABC analysis was applied to the sum of settings in which each of the 110 genes was involved (column sums in the matrix plot in Fig. 4, top). This assigned 22 genes to ABC sets 'A' or 'B' (Table 2) that were included in further analysis as they could be regarded as best suited to perform an evidence-based functional genomics analysis of chronic central sensitization to pain, considering the difficulty of separation between a modulation of the pain-causing disease from a modulation of the perception and processing of chronic nociceptive input when a gene was involved only in a single clinical setting of persisting pain.

ORA identified 64 GO terms associated with this particular subset of genes more often than expected by chance, given the chosen p -value thresholds of 1×10^{-6} and correction for multiple testing according to Bonferroni. These terms provided a functional genomics perspective of the genes with variants

Figure 3 Result of a projection of the genes carrying variants reportedly modulating persistent pain onto a self-organizing map (SOM) of the Kohonen type (Kohonen, 1982). Following projection of the genes on the grid of neurons, based on their functional annotations in the Gene Ontology knowledge-base (Ashburner et al., 2000) ($n = 109$; one of the 110 genes in Table 1 was not annotated in the GO), the distance and density structures in the high-dimensional space were visualized using the so-called U*-matrix (Ultsch, 2003). Specifically, a trained SOM represents a topology preserving mapping of n high-dimensional data points $x_i \in D$, where D denotes the data space, onto a two-dimensional grid of neurons. A neuron n and the neurons in its neighbourhood $N(n)$ on the output grid of the SOM represent points in the data space. The sum of distances between n and $N(n)$ in the high-dimensional space, combined with the respective density probabilities, is shown on a U*-matrix as a height value (U-height) at neuron n . Large U-heights mean that there is a large gap in the data space. Low U-heights mean that the points are close to each other within the data space (Lötsch and Ultsch, 2014). The dots indicate the so-called best matching units (BMUs) of the SOM, which are those neurons whose weight vector is most similar to the input, i.e. the representation of the vector of the annotation of genes to GO terms. The BMUs are coloured according to the obtained clustering of the data space and labelled with the respective gene symbols. The cluster structure emerged from visualization of distances and density structures between neurons in the high-dimensional space by means of a U*-Matrix (Izenmann, 2009). Top: here, the genes represented by the BMUs are annotated. Bottom: 3D-display of the U-matrix in which the 'valleys', 'ridges' and 'basins' can be seen. Valleys indicate clusters of functionally similar genes based on the significant GO term annotations. The figure was created using the R software package (version 3.3.2 for Linux; <http://CRAN.R-project.org/>; R Development Core Team, 2008) using our R library 'Umatrix' (<https://cran.r-project.org/package=Umatrix>; Lötsch et al., 2018a).



shown to modulate the persisting pain sensitization phenotype in different clinical settings. Their graphical representation visualized their arrangement in the GO polyhierarchy (Fig. 5). Following application of functional abstraction (Ultsch and Lötsch, 2014), the main biological functions of the 22 gene carrying variants modulating persisting pain, were grouped around seven centres of biological activity or functional areas (Table 2; Fig. 5).

A first functional area to emerge was 'immune system process', represented in this particular gene set as an important common biological function (Fig. 5 middle left). The most significantly associated immune regulatory processes were 'regulation of immune effector processes' (GO:0002697, $p = 2.8 \times 10^{-8}$) and 'positive regulation of lymphocyte mediated immunity' (GO:0002708, $p = 2.9 \times 10^{-7}$). A second functional area centred on 'reactive oxygen species metabolic process' (GO:0072593, $p = 4.5 \times 10^{-11}$) and comprised mainly nitric oxide signalling related processes (Fig. 5 middle right) such as 'nitric oxide biosynthetic process' (GO:0006809, $p = 3.2 \times 10^{-12}$) and its regulation.

Further functional areas, however, mainly reflected processes known from previous research to contribute to pain (Lötsch et al., 2013; Ultsch et al.,

2016). This included 'response to stimulus' with several subordinate terms, such as 'response to stress' (GO:0006950, $p = 5.2 \times 10^{-7}$), comprising the reaction of the body to several challenges such as 'response to chemical' (GO:0042221, $p = 1.1 \times 10^{-10}$) and 'defense response' (GO:0006952, $p = 1.8 \times 10^{-7}$). Further subordinate areas included 'inflammatory response' (GO:0006954, $p = 2.8 \times 10^{-8}$) and 'response to other organism' (GO:0051707, $p = 2.5 \times 10^{-10}$) to which 'response to bacterium' (Fig. 5 middle left) was related. These response areas were also associated with the more general functional area 'regulation of multicellular organismal process' (GO: 51239, $p = 1.2 \times 10^{-7}$). In addition, the associated functional area 'transport', mainly comprised subordinate processes related to ion or transmitter transport such as 'regulation of ion transport' (GO:0043269, $p = 9 \times 10^{-9}$) or 'regulation of secretion by cell' (GO:1903530, $p = 4.1 \times 10^{-7}$). Finally, a functional area 'neurological system process' (GO:0050877, $p = 2.4 \times 10^{-7}$), as a more specific subordinate term to 'system process', reflected the expected involvement of nervous system related processes with pain.

The main results, i.e. the most specific functional areas pointing at immune processes and nitric oxide

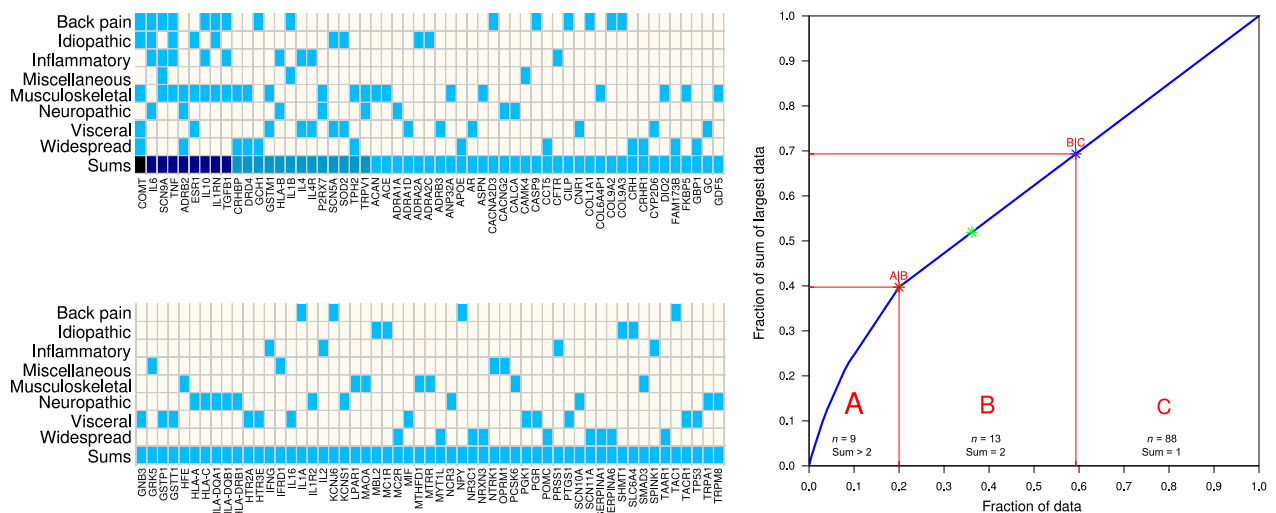


Figure 4 Number of clinical settings in which variants in the respective genes have been reported to be associated with modified phenotypes of chronic pain. Left: Matrix plot of the clinical settings (rows) versus the genes ($n = 110$, columns split in two halves, sorted for column sums in descending order). The column sums are displayed in the bottom row below the matrix. The numbers are displayed colour-coded with 0 = white, 1 = blue, >1 darker blue). Right: ABC plot of the cumulative distribution function of the sums of clinical settings in which the genes were reportedly involved (column sums in the top matrix; bottom line). The ABC set limits are indicated as red lines (for further details about an ABC analysis, see (Ultsch and Lötsch, 2015)). The figure was created using the R software package (version 3.3.2 for Linux; <http://CRAN.R-project.org/>; R Development Core Team, 2008). Specifically, the matrix plot was created using the 'heatmap.2' functions of the R package 'gplots' (Warnes G. R.; <https://cran.r-project.org/package=gplots>), with the build-in clustering of the plotting routine disabled (R switches 'Colv=FALSE, Rowv=FALSE'). The ABC curve was drawn using our R library 'ABCanalysis' (<https://cran.r-project.org/package=ABCanalysis>; Ultsch and Lötsch, 2015).

Table 2 Genes for which functional involvement of their variants has been reported in more than one clinical setting of persisting pain (number of these clinical settings given in the third column) were assigned to ABC sets 'A' or 'B' (Figure 4), comprising the most profitable information for an evidence-based functional genomics analysis of persisting pain.

Gene symbol	Gene name	Number of clinical settings of persisting pain	Functional areas						
			Response to stimulus	Immune system process	Reactive oxygen species metabolic process	Transport	Neurological system process	Regulation of multicellular organismal process	Multiorganism process
COMT	Catechol-O-methyltransferase	5	X	O	O	X	X	X	X
IL-6	Interleukin-6	4	X	X	X	X	X	X	X
SCN9A	Sodium channel, voltage-gated, type IX alpha subunit	4	X	O	O	X	X	O	O
TNF	Tumour necrosis factor	4	X	X	X	X	X	X	X
ADRB2	Adrenoceptor beta 2, surface	3	X	O	O	X	O	X	O
ESR1	Oestrogen receptor 1	3	X	O	X	O	O	X	X
IL-10	Interleukin-10	3	X	X	X	X	X	X	X
IL-1RN	Interleukin-1 receptor antagonist	3	X	X	O	X	X	X	X
TGFB1	Transforming growth factor, beta 1	3	X	X	X	X	O	X	X
CRHBP	Corticotropin-releasing hormone binding protein	2	X	O	O	X	X	X	X
DRD4	Dopamine receptor D4	2	X	O	O	X	X	X	X
GCH1	GTP cyclohydrolase 1	2	X	X	X	O	X	O	X
GSTM1	Glutathione S-transferase mu 1	2	X	O	O	O	O	O	O
HLA-B	Major histocompatibility complex, class I, B	2	X	X	O	O	O	X	X
IL-1B	Interleukin-1, beta	2	X	X	X	X	O	X	X
IL-4	Interleukin-4	2	X	X	X	X	O	X	X
IL-4R	Interleukin-4 receptor	2	X	X	O	X	O	X	X
P2RX7	Purinergic receptor P2X, ligand gated ion channel, 7	2	X	X	X	X	X	X	X
SCN5A	Sodium channel, voltage-gated, type V alpha subunit	2	X	O	O	X	X	X	O
SOD2	Superoxide dismutase 2, mitochondrial	2	X	X	X	O	X	O	X
TPH2	Tryptophan hydroxylase 2	2	X	O	O	O	O	O	O
TRPV1	Transient receptor potential cation channel, subfamily V, member 1	2	X	O	O	X	X	O	O
Sum			22	12	10	16	13	16	16

The right part of the table displays the functional areas (Figure 5) or groups of biological functions in which the gene set is involved, together with their association with each gene (X = yes, O = No). The precise definition of the GO terms can be obtained using AmiGO search tool for GO at <http://amigo.geneontology.org/> (Carbon et al., 2009).

signalling as key biological processes involved in persisting pain across more than a single clinical setting, prevailed when narrowing the gene set further on those that are also listed in the PainGenes database (Fig. 6). As the set intersection between the 22

genes identified above and the genes of the PainGenes database included only $n = 13$ genes, the ORA applying the same statistical thresholds resulted in fewer additional significant GO terms (Fig. 6).

4. Discussion

The present analysis used empirical evidence for functional human genetic variants to approach the genetic architecture of persisting pain. Although the evidence was collected from separate studies, its combination permitted a limited genomewide association analysis of the trait. Methods for data mining and machine-learned knowledge discovery were applied to publicly available databases in order to relate knowledge, acquired in the context of clinical pain research, on genes that modulate the phenotype of persisting pain with data on the biological functions of these human genes acquired in any context, without restriction to pain research (Lötsch et al., 2013).

The initial analysis of the whole set of genes showing positive results from clinical pain association studies indicated that, apart from a minority of genes that could be topically grouped such as interleukin or histocompatibility complex-related genes, most genes displayed very heterogeneous functions and the analysis did not illuminate the pathophysiology of persisting pain beyond the functions of the single genes. This was probably due to the fact that data were drawn mainly from candidate gene or GWAS approaches. This selection probably introduced a research bias by (1) addressing genetic modulators in the context of the underlying disease and (2) including a selection of genetic markers that mimic other successful reports of comparable studies.

The situation became clearer when the analytical focus was narrowed to genetic modulations consistently observed across several clinical conditions with potential underlying painful diseases. This reduced the analytical bias generated by genetic modulations responsible for a specific pain-causing disease and is in keeping with the contemporary approach to persisting pain as a distinct condition of central sensitization to pain and not merely a symptom of another underlying chronic disease. Consequently, it would be expected that the trait is modulated by specific genes which should be reflected by observations on its modulation in clinical research. The mathematically precise calculation provided by the ABC analysis (Ultsch and Lötsch, 2015), developed in order to select the most promising or profitable items from a larger set of items, resulted in identification of a set of 22 genes which could be then be assessed in a computational functional genomics analysis of persisting pain.

A major finding of this analysis of available evidence on genetic modulation of persisting pain was the particular importance of two groups of biological processes indicating involvement (1) of the immune

system and of (2) nitric oxide signalling in persisting pain. Involvement of both processes is biologically highly plausible; however, their emergence as major process groups from a functional genomics analysis of data from clinical genetic research on persisting pain was not anticipated. Specific roles for the present subset of 22 genes, with repeated evidence for involvement in persisting pain, were exhibited by the 12 genes annotated as 'immune system process' (Table 1). This subset included interleukin (*IL-1B*, *IL-4*, *IL-6*, *IL-10*) (Dinarello, 1994; Choi and Reiser, 1998; Mocellin et al., 2004; Nemeth et al., 2004), interleukin receptor (*IL-1RN*, *IL-4R*) (Bittar and Bittar, 1996) and histocompatibility complex-related (*HLA-B*) genes (Dupont and Ceppellini, 1989), which have been shown to be involved in immunological mechanisms of pain (Sato et al., 2002; de Rooij et al., 2009). This is also supported by published evidence for the further genes in this list, such as, *TNF* (Vassalli, 1992; Franchimont et al., 1999), *TGFB1* (Li et al., 2006), *GCH1* (Schott et al., 1993), *P2RX7* (Schwartz et al., 2009) and *SOD2* (Wells et al., 2003). The second major process group emerging from the functional genomics analysis of the key evidence for genetic modulation of clinical persisting pain was nitric oxide signalling, in particular metabolic processes, summarized in this context under the GO term 'reactive oxygen species metabolic process' which includes the genes *IL-6* (Deakin et al., 1995), *TNF* (Deakin et al., 1995; Katusic et al., 1998), *ESR1* (Clapauch et al., 2014), *IL-10* (Cattaruzza et al., 2003), *TGFB1* (Saura et al., 2005), *GCH1* (Katusic et al., 1998; Zhang et al., 2007), *IL-1B* (Katusic et al., 1998), *IL-4* (Coccia et al., 2000), *P2RX7* (Gendron et al., 2003), *SOD2* (Fridovich, 1978). It is widely accepted that nitric oxide (NO) is critically involved in persisting pain (Chung, 2004). It has been shown that NO is produced in the spinal dorsal horn neurons in response to extensive nociceptive inputs and then it diffuses out and increases neurotransmitter release from primary afferent terminals, thereby contributing to central sensitization and persisting pain (Lin et al., 1999). Recent findings seem to indicate that not only NO is a mediator of persisting pain that accompanies inflammation, other reactive oxygen species like superoxide (SO) might also participate in persisting pain (Schwartz et al., 2008). Kim and colleagues found that NO and SO contribute to persisting pain via two separate and independent pathways and a recent study has shown that capsaicin-induced hyperalgesia is a consequence of superoxide build-up in spinal dorsal horn neurons. Superoxide dismutase (SOD-2) encoded by gene *SOD2* is a major

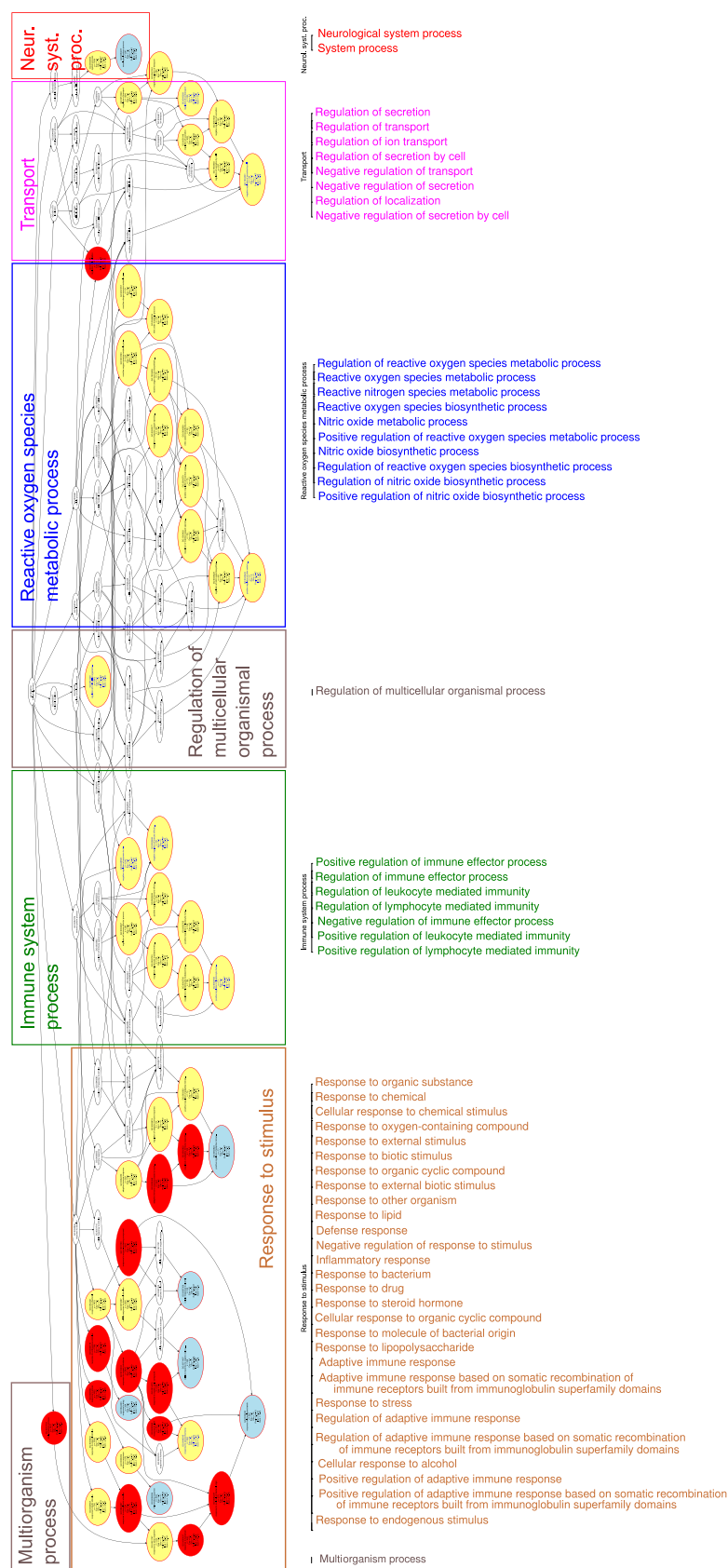


Figure 5 Results of an overrepresentation analysis (ORA; p -value threshold, $t_p = 1 \times 10^{-6}$ and Bonferroni α correction) of 22 genes (Table 2) carrying variants reported to modulate the persisting pain phenotypes in different clinical settings (ABC sets 'A' or 'B') versus all human genes. A top-down representation of the annotations (GO terms) is shown representing a systems-biology perspective of the biological processes modulated by this set of genes. Each ellipse represents a GO term. The graphical representation follows the standard of the GO knowledge base, where GO terms are related to each other by 'is-a', 'part-of', and 'regulates' relationships forming a branching polyhierarchy organized in a directed acyclic graph (DAG; Thulasiraman and Swamy, 1992). Top: Significant GO terms are shown as coloured ellipses with the number of member genes, the number of expected genes by chance and the significance of the deviation in the observed from the expected number of genes indicated. The biological processes in which the present $n = 22$ genes are involved can be summarized by seven primary 'functional areas' or headlines presenting particular aspects (taxonomies) of the complete polyhierarchy at maximum coverage, precision, informational value and conciseness (Ultsch and Löttsch, 2014). The ellipses are colour-coded using yellow for a 'headline', i.e. a GO term that by its location in the polyhierarchy may serve as headlines for a branch of the hierarchy, red for significantly overrepresented terms and white for non-significant terms that need to be displayed to preserve the polyhierarchical structure of the DAG. Blue vertices or blue labels are the most specific terms (leaves of the DAG) at the end of a taxonomy (branch) in the polyhierarchy. Bottom: The GO terms (biological processes) taken from the functional areas, shown above in the DAG, are shown with larger fonts for better readability. The figure was created using the R software package (version 3.3.2 for Linux; <http://CRAN.R-project.org/>; R Development Core Team, 2008) and our R package 'dbtORA' (<https://github.com/IME-TMP-FFM/dbtORA>; Lippmann et al., 2018).

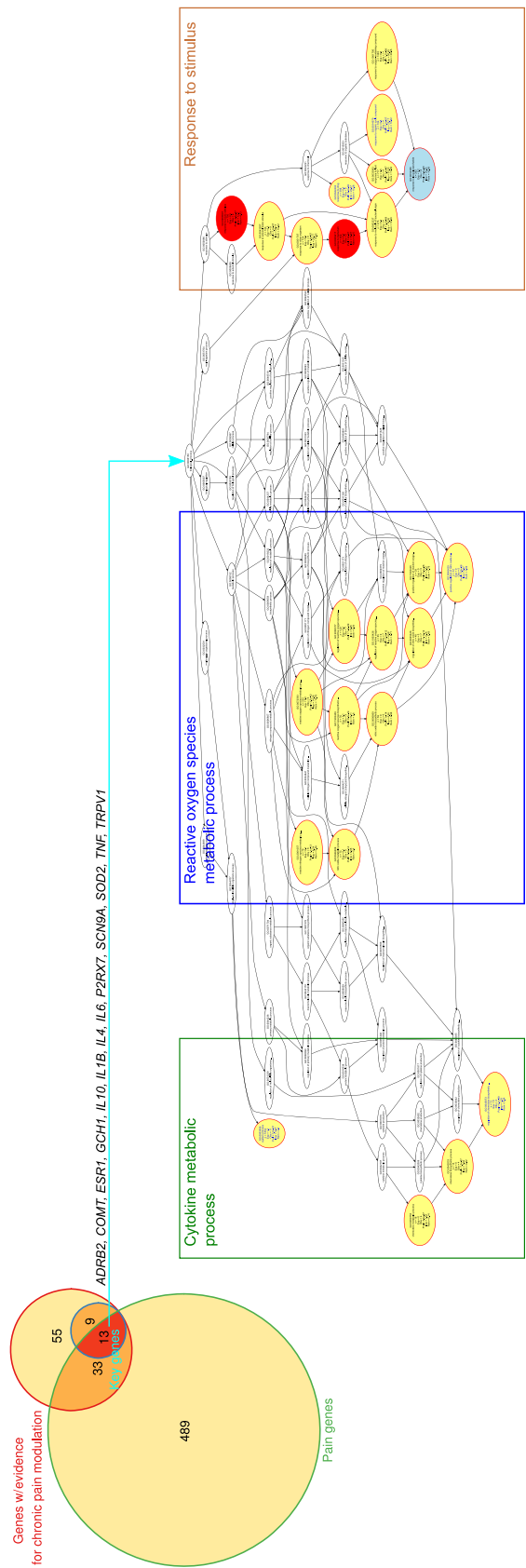


Figure 6 Overrepresentation analysis (ORA) of 13 genes (Table 2), versus all human genes, that (1) carry variants reported to modulate the persisting pain phenotypes, (2) belonging to the subset of 22 of these genes that are supported by evidence that this modulation applies to several clinical settings (ABC sets 'A' or 'B') and (3) were also part of a set of 535 genes, based on current pain research data, that can be considered, with their products, to be involved in the pathophysiology of pain (Ultsch et al., 2016). Left part: The Venn diagram (Venn, 1880) shows the set intersections on which this selection of 13 genes is based. Right part: Top-down representation of the annotations (GO terms) representing the systems-biology perspective of the biological processes modulated by this set of genes organized in a branching polyhierarchy forming a directed acyclic graph (DAG; Thulasiraman and Swamy, 1992). The figure represents the results of an overrepresentation analysis with parameters for p -value threshold, $t_p = 1 \times 10^{-6}$ and Bonferroni α correction. The biological processes in which the $n = 13$ genes are involved can be summarized by three primary 'functional areas' or headlines presenting particular aspects (taxonomies) of the complete polyhierarchy at maximum coverage, precision, informational value and conciseness (Ultsch and Löttsch, 2014). The ellipses are colour-coded using yellow for a 'headline'; i.e. a GO term that by its location in the polyhierarchy may serve as headlines for a branch of the hierarchy, red for significantly overrepresented terms located in the polyhierarchy below a headline and white for nonsignificant terms that need to be displayed to preserve the polyhierarchical structure of the DAG. Blue vertices or blue labels are the most specific terms (leaves of the DAG) at the end of a taxonomy (branch) in the polyhierarchy. Bottom: The GO terms (biological processes) taken from the functional areas, shown above in the DAG, are shown with larger fonts for better readability. The figure was created using the R software package (version 3.3.2 for Linux; <http://CRAN.R-project.org/>; R Development Core Team, 2008) and our R package 'dbtORA' (<https://github.com/IME-TMP-FFM/dbtORA>; Lippmann et al., 2018). The Venn diagram was drawn using the R library 'Vennable' (Swinton J., https://r-forge.r-project.org/R/?group_id=474).

determinant suggesting a therapeutic potential for the manipulation of spinal SOD-2 activity in pain conditions (Schwartz et al., 2009).

The role of cytokines was further highlighted by further restricting the gene subset most consistently related to persisting pain by identifying the appearance of the relevant gene in an independently created list of 535 so-called pain genes (Ultsch et al., 2016). These were genes relevant to pain listed by several sources: mainly the Pain Genes Database (<http://www.jblsdesign.com/jmogil/enter.html> (Lacroix-Fralish et al., 2007). The overlap of the set of the 22 genes, with repeated records of variants that modulate the clinical phenotype of persisting pain, with the 535 'pain genes' comprised 13 genes (*ADRB2*, *COMT*, *ESR1*, *GCH1*, *IL-10*, *IL-1B*, *IL-4*, *IL-6*, *P2RX7*, *SCN9A*, *SOD2*, *TNF*, *TRPV1*). After performing a similar ORA (Fig. 6) as that described above, these 13 genes were found to be mainly involved in cytokine production, covered by the significant GO terms 'chemokine metabolic process' (GO:0050755, $p = 3.3 \times 10^{-7}$), 'nitrogen compound metabolic process' and again 'response to stimulus', the most significant term being 'nitric oxide biosynthetic process' (GO:0006809, $p = 1 \times 10^{-14}$).

The involvement of the immune system in persisting pain is plausible from a biological perspective. One of the sites of interaction of the immune system with persisting pain has been identified as neuroimmune crosstalk at the glial–opioid interface (Tian et al., 2012). Previous research has shown that glial and immune cells, including astrocytes, microglia/macrophages, as well as T lymphocytes, are key cells activated during persisting pain, which contribute to pain persistence (Calvo et al., 2012; von Hehn et al., 2012). A role for local production of cytokines in the central nervous system during inflammatory conditions associated with persisting pain, such as rheumatoid arthritis (Lampa et al., 2012) or fibromyalgia (Kadetoff et al., 2012), as well as evidence for central nervous system sensitization by cytokines (Aden et al., 2010) also suggests such an immune system interaction with persisting pain states. Indeed, it has recently been suggested that the predominance of pain sensitization during chronic diseases in women is closely linked to the effects of female sex hormones on the neuroimmune system (Rosen et al., 2017).

The present results clearly support the modulation of neuroimmune system processes as a promising strategy in the development of novel analgesic drugs against persisting pain. This may be possible along several lines. For example, involvement of glial cells in opiate actions has been shown recently (Chen et al.,

2010; Boue et al., 2012). Consequently, the elucidation of pain- and opioid-induced mechanisms at the level of glial and immune cells could lead to improvement of pain management. In an animal model, ibudilast, a nonselective phosphodiesterase-inhibiting, anti-inflammatory drug that also blocks glial activation probably via antagonism at the Toll-like receptor 4 (Jia et al., 2012), restored morphine-induced antinociception following tolerance development (Liljus et al., 2009). Similarly, minocycline, a tetracycline that inhibits microglial activation and proliferation, also seems to attenuate morphine tolerance in mouse models of neuropathic pain (Chen et al., 2010). Hence, increasing evidence points towards the immune system as a potential source of future targets for analgesic drugs directed against persisting pain.

By gathering the relevant reports, the present analysis centres on the current evidence about a genetic modulation of persisting pain on a gene level, without going into the details of single nucleotide polymorphism level as usually applied in review papers. However, the approach was centred on machine-learned knowledge discovery from the gathered evidence and was based on published evidence gathered from studies in which the authors had used a candidate gene approach or had performed a GWAS without a gene-specific hypothesis. Therefore, the present analysis implies a research bias given by the original hypotheses or on the inclusion of frequently addressed genetic variants in the analysed studies. The question addressed with the present analysis was about the greater functional perspective emerging from successful clinical studies of the genetic modulation of persisting pain. Importantly, while the analysis is biased with respect to the gene selection, made by the authors of the included studies, its results are not biased for a particular functional genomics perspective as this had not been an gene inclusion criterion in the analysed studies. Nevertheless, the present selection of repeatedly reported associations implies an advantage of frequently included genes that have been attracted research interest through the last several years, such as *OPRM1*, *GCH1* or *COMT*.

With the caution advised by the implicit research bias regarding the gene selection, the results of the present analysis were (1) unexpected considering that hypotheses about the involvement of immune system processes or of nitric oxide signalling were not preformulated for the present analysis and (2) biologically plausible and completely compatible with current research activities on persisting pain in the light of increasingly acknowledgement of an involvement of immune processes that has attracted

concerted research activities (Kringel and Lötsch, 2015), including that on the role of the glial–opioid interface in persisting pain (<http://gloria.helsinki.fi>). Hence, although the evidence was generated in separate studies, the combination of positive findings of a genetic modulation of persisting pain allows a limited yet valid genomewide analysis of the trait, providing a more comprehensive picture of functional background of persisting pain from genomics perspective than associations of single genotypes.

5. Conclusions

While many studies have focused on particular genes, the present analysis pursued the question whether their combined results may provide more complex insights into the pathophysiology of persisting pain in humans. In keeping with the contemporary trend towards ‘big data’ analysis in biomedical research, the current empirical data on modulation of persisting pain via human genetic polymorphisms have been subjected here to a computational functional genomics analysis. This evidence was then analysed for emergent, principal pathophysiological processes that characterize persisting sensitization to pain. Analysis of 110 unique genes, with variants that have been reported to modulate the clinical phenotype of persisting pain, led to the selection of functionally heterogeneous genes. By focusing on genes that have been repeatedly associated with modulation of persisting pain phenotypes in several clinical settings, a clearer picture emerged of the main processes identified by current human genetics research on persisting pain. A mathematically supported, precise selection of a subset of genes was possible using a computational functional genomics approach. On this basis, including the research bias of current clinical genetic association studies, the evidence gathered so far points to the view of persisting pain as a trait resulting from alterations in the immune system and/or in nitric oxide signalling, a concept that is biologically highly plausible and agrees with other lines of pain research. While analysing existing evidence and therefore limited to previously shown functional pathways, the present computational functional genomics-based approach provides a computational systems-biology perspective on chronic sensitization to pain by summarizing the empirical evidence gathered in many separate studies. Moreover, human genetic research on persisting pain emphasizes the immune system as a potential source of important future targets for analgesic drugs directed against persisting pain and demonstrates

that contemporary machine-learned methods offer innovative approaches to knowledge discovery from previous evidence.

Author contributions

JL, DK and EK conceived and designed the analysis. JL, CL and DK analysed the data. JL, DK, EK and MJP wrote the article. JL, AU, DK and EK involved in discussion of methods and results.

References

- Aden, U., Favrais, G., Plaisant, F., Winerdal, M., Felderhoff-Mueser, U., Lampa, J., Lelievre, V., Gressens, P. (2010). Systemic inflammation sensitizes the neonatal brain to excitotoxicity through a pro-/anti-inflammatory imbalance: Key role of TNF α pathway and protection by etanercept. *Brain Behav Immun* 24, 747–758.
- Ala-Kokko, L. (2002). Genetic risk factors for lumbar disc disease. *Ann Med* 34, 42–47.
- Alizadeh, B.Z., Njajou, O.T., Hazes, J.M., Hofman, A., Slagboom, P.E., Pols, H.A., van Duijn, C.M. (2007). The H63D variant in the HFE gene predisposes to arthralgia, chondrocalcinosis and osteoarthritis. *Ann Rheum Dis* 66, 1436–1442.
- Aneiros-Guerrero, A., Lendinez, A.M., Palomares, A.R., Perez-Nevot, B., Aguado, L., Mayor-Olea, A., Ruiz-Galdon, M., Reyes-Engel, A. (2011). Genetic polymorphisms in folate pathway enzymes, DRD4 and GSTM1 are related to temporomandibular disorder. *BMC Med Genet* 12, 75.
- Arisan, E.D., Arisan, S., Kiremit, M.C., Tigli, H., Caskurlu, T., Palavan-Unsal, N., Ergenekon, E. (2006). Manganese superoxide dismutase polymorphism in chronic pelvic pain syndrome patients. *Prostate Cancer Prostatic Dis* 9, 426–431.
- Arisawa, T., Tahara, T., Shibata, T., Nagasaka, M., Nakamura, M. et al. (2007). Genetic polymorphisms of molecules associated with inflammation and immune response in Japanese subjects with functional dyspepsia. *Int J Mol Med* 20, 717–723.
- Arisawa, T., Tahara, T., Shibata, T., Nagasaka, M., Nakamura, M. et al. (2008). Genetic polymorphisms of cyclooxygenase-1 (COX-1) are associated with functional dyspepsia in Japanese women. *J Womens Health (Larchmt)* 17, 1039–1043.
- Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H. et al. (2000). Gene ontology: Tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* 25, 25–29.
- Attur, M., Wang, H.Y., Kraus, V.B., Bukowski, J.F., Aziz, N. et al. (2010). Radiographic severity of knee osteoarthritis is conditional on interleukin 1 receptor antagonist gene variations. *Ann Rheum Dis* 69, 856–861.
- Babula, O., Danielsson, I., Sjöberg, I., Ledger, W.J., Witkin, S.S. (2004). Altered distribution of mannose-binding lectin alleles at exon 1 codon 54 in women with vulvar vestibulitis syndrome. *Am J Obstet Gynecol* 191, 762–766.
- Backes, C., Keller, A., Kuentzer, J., Kneissl, B., Comtesse, N. et al. (2007). GeneTrail-advanced gene set enrichment analysis. *Nucleic Acids Res* 35, W186–W192.
- Binder, A., May, D., Baron, R., Maier, C., Tölle, T.R. et al. (2011). Transient receptor potential channel polymorphisms are associated with the somatosensory function in neuropathic pain patients. *PLoS ONE* 6, e17387.
- Bittar, E.E., Bittar, N. (1996). *Immunobiology* (Greenwich, Conn.: JAI Press).
- Blanco, I., Arbesú, D., Kassam, D.A., de Serres, F.J., Fernández-Bustillo, E., Rodríguez, C. (2006). Alpha1-antitrypsin polymorphism in fibromyalgia syndrome patients from the Asturias province in Northern Spain: A significantly higher prevalence of the Pi*Z deficiency allele in patients than in the general population. *J Musculoskelet Pain* 14, 5–12.

- Bortsov, A.V., Smith, J.E., Diatchenko, L., Soward, A.C., Ulirsch, J.C. et al. (2013). Polymorphisms in the glucocorticoid receptor co-chaperone FKBP5 predict persistent musculoskeletal pain after traumatic stress exposure. *Pain* 154, 1419–1426.
- Boue, J., Blanpied, C., Djata-Cabral, M., Pelletier, L., Vergnolle, N., Dietrich, G. (2012). Immune conditions associated with CD4 + T effector-induced opioid release and analgesia. *Pain* 153, 485–493.
- Breivik, H., Collett, B., Ventafridda, V., Cohen, R., Gallacher, D. (2006). Survey of chronic pain in Europe: Prevalence, impact on daily life, and treatment. *Eur J Pain* 10, 287–333.
- Bruehl, S., Denton, J.S., Loneragan, D., Koran, M.E., Chont, M. et al. (2013). Associations between KCNJ6 (GIRK2) gene polymorphisms and pain-related phenotypes. *Pain* 154, 2853–2859.
- Buskila, D., Cohen, H., Neumann, L., Ebstein, R.P. (2004). An association between fibromyalgia and the dopamine D4 receptor exon III repeat polymorphism and relationship to novelty seeking personality traits. *Mol Psychiatry* 9, 730–731.
- Calvo, M., Dawes, J.M., Bennett, D.L. (2012). The role of the immune system in the generation of neuropathic pain. *Lancet Neurol* 11, 629–642.
- Camon, E., Magrane, M., Barrell, D., Binns, D., Fleischmann, W. et al. (2003). The Gene Ontology Annotation (GOA) project: Implementation of GO in SWISS-PROT, TrEMBL, and InterPro. *Genome Res* 13, 662–672.
- Camon, E., Magrane, M., Barrell, D., Lee, V., Dimmer, E. et al. (2004). The Gene Ontology Annotation (GOA) Database: Sharing knowledge in Uniprot with Gene Ontology. *Nucleic Acids Res* 32, D262–D266.
- Carbon, S., Ireland, A., Mungall, C.J., Shu, S., Marshall, B., Lewis, S., Ami, G.O.H., Web Presence Working Group (2009). AmiGO: Online access to ontology and annotation data. *Bioinformatics* 25, 288–289.
- Cattaruzza, M., Slodowski, W., Stojakovic, M., Krzesz, R., Hecker, M. (2003). Interleukin-10 induction of nitric-oxide synthase expression attenuates CD40-mediated interleukin-12 synthesis in human endothelial cells. *J Biol Chem* 278, 37874–37880.
- Chen, S., Hui, H., Zhang, D., Xue, Y. (2010). The combination of morphine and minocycline may be a good treatment for intractable post-herpetic neuralgia. *Med Hypotheses* 75, 663–665.
- Cheng, K.I., Lin, S.R., Chang, L.L., Wang, J.Y., Lai, C.S. (2010). Association of the functional A118G polymorphism of OPRM1 in diabetic patients with foot ulcer pain. *J Diabetes Complications* 24, 102–108.
- Choi, P., Reiser, H. (1998). IL-4: Role in disease and regulation of production. *Clin Exp Immunol* 113, 317–319.
- Chung, J.M. (2004). The role of reactive oxygen species (ROS) in persistent pain. *Mol Interv* 4, 248–250.
- Clapauch, R., Mourao, A.F., Mecnas, A.S., Maranhao, P.A., Rossini, A., Bouskela, E. (2014). Endothelial function and insulin resistance in early postmenopausal women with cardiovascular risk factors: Importance of ESR1 and NOS3 polymorphisms. *PLoS ONE* 9, e103444.
- Coccia, E.M., Stellacci, E., Marziali, G., Weiss, G., Battistini, A. (2000). IFN-gamma and IL-4 differently regulate inducible NO synthase gene expression through IRF-1 modulation. *Int Immunol* 12, 977–985.
- Cohen, H., Neumann, L., Glazer, Y., Ebstein, R.P., Buskila, D. (2009). The relationship between a common catechol-O-methyltransferase (COMT) polymorphism val(158) met and fibromyalgia. *Clin Exp Rheumatol* 27, S51–S56.
- Costigan, M., Belfer, I., Griffin, R.S., Dai, F., Barrett, L.B. et al. (2010). Multiple chronic pain states are associated with a common amino acid-changing allele in KCNS1. *Brain* 133, 2519–2527.
- Dai, F., Belfer, I., Schwartz, C.E., Banco, R., Martha, J.F. et al. (2010). Association of catechol-O-methyltransferase genetic variants with outcome in patients undergoing surgical treatment for lumbar degenerative disc disease. *Spine J* 10, 949–957.
- De Carvalho, C.V., Nogueira-De-Souza, N.C., Costa, A.M., Baracat, E.C., Girao, M.J., D'Amora, P., Schor, E., da Silva, I.D. (2007). Genetic polymorphisms of cytochrome P450c17alpha (CYP17) and progesterone receptor genes (PROGINS) in the assessment of endometriosis risk. *Gynecol Endocrinol* 23, 29–33.
- Deakin, A.M., Payne, A.N., Whittle, B.J., Moncada, S. (1995). The modulation of IL-6 and TNF-alpha release by nitric oxide following stimulation of J774 cells with LPS and IFN-gamma. *Cytokine* 7, 408–416.
- Derry, S., Gill, D., Phillips, T., Moore, R.A. (2012). Milnacipran for neuropathic pain and fibromyalgia in adults. *Cochrane Database Syst Rev* 3, CD008244.
- Dhar, V. (2013). Data science and prediction. *Commun ACM* 56, 64–73.
- Diatchenko, L., Anderson, A.D., Slade, G.D., Fillingim, R.B., Shabalina, S.A. et al. (2006). Three major haplotypes of the beta2 adrenergic receptor define psychological profile, blood pressure, and the risk for development of a common musculoskeletal pain disorder. *Am J Med Genet B Neuropsychiatr Genet* 141B, 449–462.
- Dinarello, C.A. (1994). The biological properties of interleukin-1. *Eur Cytokine Netw* 5, 517–531.
- Docampo, E., Escaramis, G., Gratacos, M., Villatoro, S., Puig, A. et al. (2014). Genome-wide analysis of single nucleotide polymorphisms and copy number variants in fibromyalgia suggest a role for the central nervous system. *Pain* 155, 1102–1109.
- Doehring, A., Freynhagen, R., Griessinger, N., Zimmermann, M., Sittl, R., Hentig, N., Geisslinger, G., Lötsch, J. (2009). Cross-sectional assessment of the consequences of a GTP cyclohydrolase 1 haplotype for specialized tertiary outpatient pain care. *Clin J Pain* 25, 781–785.
- Dupont, B., Ceppellini, R. (1989). *Immunobiology of HLA* (New York: Springer-Verlag).
- Edwards, R.R. (2006). Genetic predictors of acute and chronic pain. *Curr Rheumatol Rep* 8, 411–417.
- Elliott, A.M., Smith, B.H., Penny, K.I., Smith, W.C., Chambers, W.A. (1999). The epidemiology of chronic pain in the community. *Lancet* 354, 1248–1252.
- Faber, C.G., Lauria, G., Merkies, I.S., Cheng, X., Han, C. et al. (2012). Gain-of-function Nav1.8 mutations in painful neuropathy. *Proc Natl Acad Sci USA* 109, 19444–19449.
- Faserl, K., Golderer, G., Kremser, L., Lindner, H., Sarg, B., Wildt, L., Seeber, B. (2011). Polymorphism in vitamin D-binding protein as a genetic risk factor in the pathogenesis of endometriosis. *J Clin Endocrinol Metab* 96, E233–E241.
- Fisher, R.A. (1922). On the interpretation of chi square from contingency tables, and the calculation of P. *J Roy Stat Soc* 85, 87–94.
- Foster, D.C., Sazenski, T.M., Stodgell, C.J. (2004). Impact of genetic variation in interleukin-1 receptor antagonist and melanocortin-1 receptor genes on vulvar vestibulitis syndrome. *J Reprod Med* 49, 503–509.
- Franchimont, D., Martens, H., Hagelstein, M.T., Louis, E., Dewe, W., Chrousos, G.P., Belaiche, J., Geenen, V. (1999). Tumor necrosis factor alpha decreases, and interleukin-10 increases, the sensitivity of human monocytes to dexamethasone: Potential regulation of the glucocorticoid receptor. *J Clin Endocrinol Metab* 84, 2834–2839.
- Fridovich, I. (1978). The biology of oxygen radicals. *Science* 201, 875–880.
- Gan, X.L., Lin, Y.H., Zhang, Y., Yu, T.H., Hu, L.N. (2010). Association of an interleukin-16 gene polymorphism with the risk and pain phenotype of endometriosis. *DNA Cell Biol* 29, 663–667.
- Gendron, F.P., Chalimoniuk, M., Strosznajder, J., Shen, S., Gonzalez, F.A., Weisman, G.A., Sun, G.Y. (2003). P2X7 nucleotide receptor activation enhances IFN gamma-induced type II nitric oxide synthase activity in BV-2 microglial cells. *J Neurochem* 87, 344–352.
- Gibson, G., Muse, S.V. (2009). *A Primer of Genome Science* (Sunderland, Massachusetts: Sinauer Associates).
- Godinova, A.M. (1965). Genetic analysis of migraine. *Zh Nevropatol Psikiatr Im S S Korsakova* 65, 1132–1138.
- Govindan, S., Shaik, N.A., Vedicherla, B., Kodati, V., Rao, K.P., Hasan, Q. (2009). Estrogen receptor-alpha gene (T/C) Pvu II polymorphism in endometriosis and uterine fibroids. *Dis Markers* 26, 149–154.
- Gullo, L., Tabacchi, P.L., Corazza, G.R., Calanca, F., Campione, O., Labo, G. (1982). HLA-B13 and chronic calcific pancreatitis. *Dig Dis Sci* 27, 214–216.
- Guo, T.M., Liu, M., Zhang, Y.G., Guo, W.T., Wu, S.X. (2011). Association between Caspase-9 promoter region polymorphisms and discogenic low back pain. *Connect Tissue Res* 52, 133–138.
- Gursoy, S., Erdal, E., Sezgin, M., Barlas, I.O., Aydeniz, A., Alasehirli, B., Sahin, G. (2008). Which genotype of MAO gene that the patients have are likely to be most susceptible to the symptoms of fibromyalgia? *Rheumatol Int* 28, 307–311.

- van Hecke, O., Torrance, N., Smith, B.H. (2013). Chronic pain epidemiology and its clinical relevance. *Br J Anaesth* 111, 13–18.
- von Hehn, C.A., Baron, R., Woolf, C.J. (2012). Deconstructing the neuropathic pain phenotype to reveal neural mechanisms. *Neuron* 73, 638–652.
- Herken, H., Erdal, E., Mutlu, N., Barlas, O., Cataloluk, O., Oz, F., Guray, E. (2001). Possible association of temporomandibular joint pain and dysfunction with a polymorphism in the serotonin transporter gene. *Am J Orthod Dentofacial Orthop* 120, 308–313.
- Herlyn, P., Muller-Hilke, B., Wendt, M., Hecker, M., Mittlmeier, T., Gradl, G. (2010). Frequencies of polymorphisms in cytokines, neurotransmitters and adrenergic receptors in patients with complex regional pain syndrome type I after distal radial fracture. *Clin J Pain* 26, 175–181.
- Hochberg, Y. (1988). A sharper bonferroni procedure for multiple tests of significance. *Biometrika* 75, 800–802.
- Hocking, L.J., Smith, B.H., Jones, G.T., Reid, D.M., Strachan, D.P., Macfarlane, G.J. (2010). Genetic variation in the beta2-adrenergic receptor but not catecholamine-O-methyltransferase predisposes to chronic pain: Results from the 1958 British Birth Cohort Study. *Pain* 149, 143–151.
- Holliday, K.L., Nicholl, B.I., Macfarlane, G.J., Thomson, W., Davies, K.A., McBeth, J. (2010). Genetic variation in the hypothalamic-pituitary-adrenal stress axis influences susceptibility to musculoskeletal pain: Results from the EPIFUND study. *Ann Rheum Dis* 69, 556–560.
- Hooten, W.M., Hartman, W.R., Black, J.L., Laures, H.J., Walker, D.L. (2013). Associations between serotonin transporter gene polymorphisms and heat pain perception in adults with chronic pain. *BMC Med Genet* 14, 78.
- Izenmann, A. (2009). *Modern Multivariate Statistical Techniques* (Berlin: Springer).
- Jeremias, J., Ledger, W.J., Witkin, S.S. (2000). Interleukin 1 receptor antagonist gene polymorphism in women with vulvar vestibulitis. *Am J Obstet Gynecol* 182, 283–285.
- Jia, Z.J., Wu, F.X., Huang, Q.H., Liu, J.M. (2012). Toll-like receptor 4: The potential therapeutic target for neuropathic pain. *Zhongguo Yi Xue Ke Xue Yuan Xue Bao* 34, 168–173.
- Juran, J.M. (1975). The non-Pareto principle; Mea culpa. *Quality Progress* 8, 8–9.
- Kadetoff, D., Lampa, J., Westman, M., Andersson, M., Kosek, E. (2012). Evidence of central inflammation in fibromyalgia-increased cerebrospinal fluid interleukin-8 levels. *J Neuroimmunol* 242, 33–38.
- Kales, S.N., Linos, A., Chatzis, C., Sai, Y., Halla, M., Nasioulas, G., Christiani, D.C. (2004). The role of collagen IX tryptophan polymorphisms in symptomatic intervertebral disc disease in Southern European patients. *Spine (Phila Pa 1976)* 29, 1266–1270.
- Kang, S.C., Lee, D.G., Choi, J.H., Kim, S.T., Kim, Y.K., Ahn, H.J. (2007). Association between estrogen receptor polymorphism and pain susceptibility in female temporomandibular joint osteoarthritis patients. *Int J Oral Maxillofac Surg* 36, 391–394.
- Karling, P., Danielsson, A., Wikgren, M., Soderstrom, I., Del-Favero, J., Adolfsson, R., Norrback, K.F. (2011). The relationship between the val158met catechol-O-methyltransferase (COMT) polymorphism and irritable bowel syndrome. *PLoS ONE* 6, e18035.
- Katusic, Z.S., Stelter, A., Milstien, S. (1998). Cytokines stimulate GTP cyclohydrolase I gene expression in cultured human umbilical vein endothelial cells. *Arterioscler Thromb Vasc Biol* 18, 27–32.
- Kilpatrick, L.A., Labus, J.S., Coveleskie, K., Hammer, C., Rappold, G. et al. (2011). The HTR3A polymorphism c.-42C>T is associated with amygdala responsiveness in patients with irritable bowel syndrome. *Gastroenterology* 140, 1943–1951.
- Kim, H.J., Camilleri, M., Carlson, P.J., Cremonini, F., Ferber, I. et al. (2004). Association of distinct alpha(2) adrenoceptor and serotonin transporter polymorphisms with constipation and somatic symptoms in functional gastrointestinal disorders. *Gut* 53, 829–837.
- Kim, D.H., Dai, F., Belfer, I., Banco, R.J., Martha, J.F. et al. (2010a). Polymorphic variation of the guanosine triphosphate cyclohydrolase 1 gene predicts outcome in patients undergoing surgical treatment for lumbar degenerative disc disease. *Spine (Phila Pa 1976)* 35, 1909–1914.
- Kim, D.H., Lee, S.H., Kim, K.T., Yu, S.D. (2010b). Association of interleukin-1 receptor antagonist gene polymorphism with response to conservative treatment of lumbar herniated nucleus pulposus. *Spine (Phila Pa 1976)* 35, 1527–1531.
- Kirk, K.M., Doege, K.J., Hecht, J., Bellamy, N., Martin, N.G. (2003). Osteoarthritis of the hands, hips and knees in an Australian twin sample—evidence of association with the aggrecan VNTR polymorphism. *Twin Res* 6, 62–66.
- Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Biol Cybernet* 43, 59–69.
- Kringel, D., Lötsch, J. (2015). Pain research funding by the European Union Seventh Framework Programme. *Eur J Pain* 19, 595–600.
- Kunz, M., Hennig, J., Karmann, A.J., Lautenbacher, S. (2016). Relationship of 5-HTTLPR polymorphism with various factors of pain processing: Subjective experience. Motor responsiveness and catastrophizing. *PLoS ONE* 11, e0153089.
- Lacroix-Fralich, M.L., Ledoux, J.B., Mogil, J.S. (2007). The Pain Genes Database: An interactive web browser of pain-related transgenic knockout studies. *Pain* 131(3), e1–e4.
- Lampa, J., Westman, M., Kadetoff, D., Agreus, A.N., Le Maitre, E. et al. (2012). Peripheral inflammatory disease associated with centrally activated IL-1 system in humans and mice. *Proc Natl Acad Sci USA* 109, 12728–12733.
- Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C. et al. (2001). Initial sequencing and analysis of the human genome. *Nature* 409, 860–921.
- Leipold, E., Liebmann, L., Korenke, G.C., Heinrich, T., Giesselmann, S. et al. (2013). A de novo gain-of-function mutation in SCN11A causes loss of pain perception. *Nat Genet* 45, 1399–1404.
- Li, M.O., Wan, Y.Y., Sanjabi, S., Robertson, A.K., Flavell, R.A. (2006). Transforming growth factor-beta regulation of immune responses. *Annu Rev Immunol* 24, 99–146.
- Lilius, T.O., Rauhalta, P.V., Kambur, O., Kalso, E.A. (2009). Modulation of morphine-induced antinociception in acute and chronic opioid treatment by ibudilast. *Anesthesiology* 111, 1356–1364.
- Lin, Q., Palecek, J., Palecková, V., Peng, Y.B., Wu, J., Cui, M., Willis, W.D. (1999). Nitric oxide mediates the central sensitization of primate spinothalamic tract neurons. *J Neurophysiol* 81, 1075–1085.
- Lindstedt, F., Lonsdorf, T.B., Schalling, M., Kosek, E., Ingvar, M. (2011). Perception of thermal pain and the thermal grill illusion is associated with polymorphisms in the serotonin transporter gene. *PLoS ONE* 6, e17752.
- Linnstaedt, S.D., Bortsov, A.V., Soward, A.C., Swor, R., Peak, D.A. et al. (2016). CRHBP polymorphisms predict chronic pain development following motor vehicle collision. *Pain* 157, 273–279.
- Lippmann, C., Kringel, D., Ultsch, A., Lötsch, J. (2018). Computational functional genomics-based approaches in analgesic drug discovery and repurposing. *Pharmacoeconomics* 19, 783–797.
- Lötsch, J., Geisslinger, G. (2011). Pharmacogenetics of new analgesics. *Br J Pharmacol* 163, 447–460.
- Lötsch, J., Ultsch, A. (2017a). Machine learning in pain research. *Pain* 159, 623–630.
- Lötsch, J., Ultsch, A. (2014). Exploiting the structures of the U-matrix. In *Advances in Intelligent Systems and Computing*, Villmann, T., Schleif, F.-M., Kaden, M., Lange, M., eds. (Heidelberg: Springer) 248–257.
- Lötsch, J., Ultsch, A. (2016a). A machine-learned computational functional genomics-based approach to drug classification. *Eur J Clin Pharmacol* 72, 1449–1461.
- Lötsch, J., Ultsch, A. (2016b). Process pharmacology: A pharmacological data science approach to drug development and therapy. *CPT Pharmacometrics Syst Pharmacol* 5, 192–200.
- Lötsch, J. and Ultsch, A. (2017b). Random forests followed by ABC analysis as a feature selection method for machine-learning. In *Conference of the International Federation of Classification Societies (Tokyo)*, p. 170.
- Lötsch, J., Doehring, A., Mogil, J.S., Arndt, T., Geisslinger, G., Ultsch, A. (2013). Functional genomics of pain in analgesic drug development and therapy. *Pharmacol Ther* 139, 60–70.

- Lötsch, J., Lippmann, C., Kringel, D., Ultsch, A. (2017). Integrated computational analysis of genes associated with human hereditary insensitivity to pain. A drug repurposing perspective. *Front Neurosci* 10, 252.
- Lötsch, J., Lerch, F., Djaldetti, R., Tegeder, I., Ultsch, A. (2018a). Identification of disease-distinct complex biomarker patterns by means of unsupervised machine-learning using an interactive R toolbox (Umatrix). *BMC Big Data Analytics* 3, pp. 5 <https://doi.org/10.1186/s41044-018-0032-1>
- Lötsch, J., Sipilä, R., Dimova, V., Kalso, E. (2018b). Machine-learned selection of psychological questionnaire items relevant to the development of persistent pain after breast cancer surgery. *Br J Anaesth* [in press].
- Malfait, A.M., Seymour, A.B., Gao, F., Tortorella, M.D., Le Graverand-Gastineau, M.P. et al. (2012). A role for PACE4 in osteoarthritis pain: Evidence from human genetic association and null mutant phenotype. *Ann Rheum Dis* 71, 1042–1048.
- Meulenbelt, I., Min, J.L., Bos, S., Riyazi, N., Houwing-Duistermaat, J.J. et al. (2008). Identification of DIO2 as a new susceptibility locus for symptomatic osteoarthritis. *Hum Mol Genet* 17, 1867–1875.
- van Meurs, J.B., Uitterlinden, A.G., Stolk, L., Kerkhof, H.J., Hofman, A., Pols, H.A., Bierma-Zeinstra, S.M. (2009). A functional polymorphism in the catechol-O-methyltransferase gene is associated with osteoarthritis-related pain. *Arthritis Rheum* 60, 628–629.
- Midha, S., Khajuria, R., Shastri, S., Kabra, M., Garg, P.K. (2010). Idiopathic chronic pancreatitis in India: Phenotypic characterisation and strong genetic susceptibility due to SPINK1 and CFTR gene mutations. *Gut* 59, 800–807.
- Miyamoto, Y., Shi, D., Nakajima, M., Ozaki, K., Sudo, A. et al. (2008). Common variants in DVWA on chromosome 3p24.3 are associated with susceptibility to knee osteoarthritis. *Nat Genet* 40, 994–998.
- Mocellin, S., Marincola, F., Rossi, C.R., Nitti, D., Lise, M. (2004). The multifaceted relationship between IL-10 and adaptive immunity: Putting together the pieces of a puzzle. *Cytokine Growth Factor Rev* 15, 61–76.
- Mogil, J.S. (2012). Pain genetics: Past, present and future. *Trends Genet* 28, 258–266.
- Moore, R.A., Straube, S., Wiffen, P.J., Derry, S., McQuay, H.J. (2009). Pregabalin for acute and chronic pain in adults. *Cochrane Database Syst Rev* CD007076. <https://doi.org/10.1002/14651858>
- Moore, R.A., Derry, S., Aldington, D., Cole, P., Wiffen, P.J. (2012). Amitriptyline for neuropathic pain and fibromyalgia in adults. *Cochrane Database Syst Rev* 12, CD008242.
- Mototani, H., Iida, A., Nakajima, M., Furuichi, T., Miyamoto, Y. et al. (2008). A functional SNP in EDG2 increases susceptibility to knee osteoarthritis in Japanese. *Hum Mol Genet* 17, 1790–1797.
- Murphy, K.P. (2012). *Machine Learning: A Probabilistic Perspective* (Cambridge, MA, USA: The MIT Press).
- Nakamura, T., Shi, D., Tzetzis, M., Rodriguez-Lopez, J., Miyamoto, Y. et al. (2007). Meta-analysis of association between the ASPN D-repeat and osteoarthritis. *Hum Mol Genet* 16, 1676–1681.
- Neely, G.G., Hess, A., Costigan, M., Keene, A.C., Goulas, S. et al. (2010). A genome-wide Drosophila screen for heat nociception identifies alpha2delta3 as an evolutionarily conserved pain gene. *Cell* 143, 628–638.
- Nemeth, E., Rivera, S., Gabayan, V., Keller, C., Taudorf, S., Pedersen, B.K., Ganz, T. (2004). IL-6 mediates hypoferrremia of inflammation by inducing the synthesis of the iron regulatory hormone hepcidin. *J Clin Invest* 113, 1271–1276.
- Nicholl, B.I., Holliday, K.L., Macfarlane, G.J., Thomson, W., Davies, K.A. et al. (2011). Association of HTR2A polymorphisms with chronic widespread pain and the extent of musculoskeletal pain: Results from two population-based cohorts. *Arthritis Rheum* 63, 810–818.
- Nissenbaum, J., Devor, M., Seltzer, Z., Gebauer, M., Michaelis, M. et al. (2010). Susceptibility to chronic pain following nerve injury is genetically affected by CACNG2. *Genome Res* 20, 1180–1190.
- Noponen-Hietala, N., Virtanen, I., Karttunen, R., Schwenke, S., Jakkula, E. et al. (2005). Genetic variations in IL6 associate with intervertebral disc disease characterized by sciatica. *Pain* 114, 186–194.
- Oen, K., Malleson, P.N., Cabral, D.A., Rosenberg, A.M., Petty, R.E., Nickerson, P., Reed, M. (2005). Cytokine genotypes correlate with pain and radiologically defined joint damage in patients with juvenile rheumatoid arthritis. *Rheumatology (Oxford)* 44, 1115–1121.
- Oshima, T., Nakajima, S., Yokoyama, T., Toyoshima, F., Sakurai, J. et al. (2010). The G-protein beta3 subunit 825 TT genotype is associated with epigastric pain syndrome-like dyspepsia. *BMC Med Genet* 11, 13.
- Ozawa, A., Sasao, Y., Iwashita, K., Miyahara, M., Sugai, J. et al. (1999). HLA-A33 and -B44 and susceptibility to postherpetic neuralgia (PHN). *Tissue Antigens* 53, 263–268.
- Park, J.M., Choi, M.G., Cho, Y.K., Lee, I.S., Kim, S.W., Choi, K.Y., Chung, I.S. (2011). Cannabinoid receptor 1 gene polymorphism and irritable bowel syndrome in the Korean population: A hypothesis-generating study. *J Clin Gastroenterol* 45, 45–49.
- Pata, C., Erdal, E., Yazc, K., Camdeviren, H., Ozkaya, M., Ulu, O. (2004). Association of the -1438 G/A and 102 T/C polymorphism of the 5-HT2A receptor gene with irritable bowel syndrome 5-HT2A gene polymorphism in irritable bowel syndrome. *J Clin Gastroenterol* 38, 561–566.
- Peters, M.J., Broer, L., Willemen, H.L., Eiriksdoottir, G., Hocking, L.J. et al. (2013). Genome-wide association study meta-analysis of chronic widespread pain: Evidence for involvement of the 5p15.2 region. *Ann Rheum Dis* 72, 427–436.
- R Development Core Team. (2008). R: A Language and Environment for Statistical Computing.
- Reeser, J.C., Payne, E., Kitchner, T., McCarty, C.A. (2011). Apolipoprotein e4 genotype increases the risk of being diagnosed with posttraumatic fibromyalgia. *PM R* 3, 193–197.
- Reimann, F., Cox, J.J., Belfer, I., Diatchenko, L., Zaykin, D.V. et al. (2010). Pain perception is altered by a nucleotide polymorphism in SCN9A. *Proc Natl Acad Sci USA* 107, 5148–5153.
- Renner, S.P., Ekcici, A.B., Maihofner, C., Oppelt, P., Thiel, F.C. et al. (2009). Neurokinin 1 receptor gene polymorphism might be correlated with recurrence rates in endometriosis. *Gynecol Endocrinol* 25, 726–733.
- Ribeiro Junior, C.L., Arruda, J.T., Silva, C.T., Moura, K.K. (2009). Analysis of p53 codon 72 gene polymorphism in Brazilian patients with endometriosis. *Genet Mol Res* 8, 494–499.
- Ribeiro-Dasilva, M.C., Peres Line, S.R., Godoy, L., dos Santos, M.C., Arthur, M.T., Hou, W., Fillingim, R.B., Rizzatti Barbosa, C.M. (2009). Estrogen receptor-alpha polymorphisms and predisposition to TMJ disorder. *J Pain* 10, 527–533.
- Riley, D.E., Krieger, J.N. (2002). X Chromosomal short tandem repeat polymorphisms near the phosphoglycerate kinase gene in men with chronic prostatitis. *Biochim Biophys Acta* 1586, 99–107.
- Rommel, O., Kley, R.A., Dekomien, G., Epplen, J.T., Vorgerd, M., Hasenbring, M. (2006). Muscle pain in myophosphorylase deficiency (McArdle's disease): The role of gender, genotype, and pain-related coping. *Pain* 124, 295–304.
- de Rooij, A.M., Florencia Gosso, M., Haasnoot, G.W., Marinus, J., Verduijn, W., Claas, F.H., van den Maagdenberg, A.M., van Hilten, J.J. (2009). HLA-B62 and HLA-DQ8 are associated with Complex Regional Pain Syndrome with fixed dystonia. *Pain* 145, 82–85.
- Rosen, S., Ham, B., Mogil, J.S. (2017). Sex differences in neuroimmunity and pain. *J Neurosci Res* 95, 500–508.
- Saito, Y.A., Strege, P.R., Tester, D.J., Locke, G.R. III, Talley, N.J. et al. (2009). Sodium channel mutation in irritable bowel syndrome: Evidence for an ion channelopathy. *Am J Physiol Gastrointest Liver Physiol* 296, G211–G218.
- Sato, M., Ohashi, J., Tsuchiya, N., Kashiwase, K., Ishikawa, Y. et al. (2002). Association of HLA-A*3303-B*4403-DRB1*1302 haplotype, but not of TNFA promoter and NKP30 polymorphism, with postherpetic neuralgia (PHN) in the Japanese population. *Genes Immun* 3, 477–481.
- Saura, M., Zaragoza, C., Herranz, B., Griera, M., Diez-Marques, L., Rodriguez-Puyol, D., Rodriguez-Puyol, M. (2005). Nitric oxide regulates transforming growth factor-beta signaling in endothelial cells. *Circ Res* 97, 1115–1123.
- Schott, K., Gutlich, M., Ziegler, I. (1993). Induction of GTP-cyclohydrolase I mRNA expression by lectin activation and

- interferon-gamma treatment in human cells associated with the immune response. *J Cell Physiol* 156, 12–16.
- Schwartz, E.S., Lee, I., Chung, K., Chung, J.M. (2008). Oxidative stress in the spinal cord is an important contributor in capsaicin-induced mechanical secondary hyperalgesia in mice. *Pain* 138, 514–524.
- Schwartz, E.S., Kim, H.Y., Wang, J., Lee, I., Klann, E., Chung, J.M., Chung, K. (2009). Persistent pain is dependent on spinal mitochondrial antioxidant levels. *J Neurosci* 29, 159–168.
- Seki, S., Kawaguchi, Y., Chiba, K., Mikami, Y., Kizawa, H. et al. (2005). A functional SNP in CILP, encoding cartilage intermediate layer protein, is associated with susceptibility to lumbar disc disease. *Nat Genet* 37, 607–612.
- Shaik, N.A., Govindan, S., Kodati, V., Rao, K.P., Hasan, Q. (2009). Polymorphic (CAG)_n repeats in the androgen receptor gene: A risk marker for endometriosis and uterine leiomyomas. *Hematol Oncol Stem Cell Ther* 2, 289–293.
- Shatzky, S., Moses, S., Levy, J., Pinsk, V., Hershkovitz, E. et al. (2000). Congenital insensitivity to pain with anhidrosis (CIPA) in Israeli-Bedouins: Genetic heterogeneity, novel mutations in the TRKA/NGF receptor gene, clinical findings, and results of nerve conduction studies. *Am J Med Genet* 92, 353–360.
- Shoskes, D.A., Albakri, Q., Thomas, K., Cook, D. (2002). Cytokine polymorphisms in men with chronic prostatitis/chronic pelvic pain syndrome: Association with diagnosis and treatment response. *J Urol* 168, 331–335.
- Smith, S.B., Maixner, D.W., Greenspan, J.D., Dubner, R., Fillingim, R.B. et al. (2011). Potential genetic risk factors for chronic TMD: Genetic associations from the OPPERA case control study. *J Pain* 12, T92–T101.
- Smith, S.B., Maixner, D.W., Fillingim, R.B., Slade, G., Gracely, R.H. et al. (2012). Large candidate gene association study reveals genetic risk factors and therapeutic targets for fibromyalgia. *Arthritis Rheum* 64, 584–593.
- Solovieva, S., Leino-Arjas, P., Saarela, J., Luoma, K., Raininko, R., Riihimäki, H. (2004). Possible association of interleukin 1 gene locus polymorphisms with low back pain. *Pain* 109, 8–19.
- Sorge, R.E., Trang, T., Dorfman, R., Smith, S.B., Beggs, S. et al. (2012). Genetically determined P2X7 receptor pore formation regulates variability in chronic pain sensitivity. *Nat Med* 18, 595–599.
- Stephens, K., Cooper, B.A., West, C., Paul, S.M., Baggott, C.R. et al. (2014). Associations between cytokine gene variations and severe persistent breast pain in women following breast cancer surgery. *J Pain* 15, 169–180.
- Sugaya, K., Nishijima, S., Yamada, T., Miyazato, M., Hatano, T., Ogawa, Y. (2002). Molecular analysis of adrenergic receptor genes and interleukin-4/interleukin-4 receptor genes in patients with interstitial cystitis. *J Urol* 168, 2668–2671.
- Tahara, T., Arisawa, T., Shibata, T., Nakamura, M., Wang, F., Hirata, I. (2008). COMT gene val158met polymorphism in patients with dyspeptic symptoms. *Hepatogastroenterology* 55, 979–982.
- Tegeder, I., Lötsch, J. (2009). Current evidence for a modulation of low back pain by human genetic variants. *J Cell Mol Med* 13(8B), 1605–19. <https://doi.org/10.1111/j.1582-4934.2009.00703.x>. Epub 2009 Feb 17.
- Tegeder, I., Costigan, M., Griffin, R.S., Abele, A., Belfer, I. et al. (2006). GTP cyclohydrolase and tetrahydrobiopterin regulate pain sensitivity and persistence. *Nat Med* 12, 1269–1277.
- Tegeder, I., Adolph, J., Schmidt, H., Woolf, C.J., Geisslinger, G., Lötsch, J. (2008). Reduced hyperalgesia in homozygous carriers of a GTP cyclohydrolase 1 haplotype. *Eur J Pain* 12, 1069–1077.
- Thulasiraman, K., Swamy, M.N.S. (1992). *Graphs: Theory and Algorithms* (New York, NY: Wiley).
- Tian, L., Ma, L., Kaarela, T., Li, Z. (2012). Neuroimmune crosstalk in the central nervous system and its significance for neurological diseases. *J Neuroinflammation* 9, 155.
- Tilkeridis, C., Bei, T., Garantzios, S., Stratakis, C.A. (2005). Association of a COL1A1 polymorphism with lumbar disc disease in young military recruits. *J Med Genet* 42, e44.
- Ultsch, A. (2003). Maps for visualization of high-dimensional data spaces. In *WSOM* (Kyushu, Japan), pp. 225–230.
- Ultsch, A. (2005). Clustering with SOM: U*C. In *Workshop on Self-Organizing Maps (Paris)*, pp. 75–82.
- Ultsch, A., Lötsch, J. (2014). Functional abstraction as a method to discover knowledge in gene ontologies. *PLoS ONE* 9, e90191.
- Ultsch, A., Lötsch, J. (2015). Computed ABC analysis for rational selection of most informative variables in multivariate data. *PLoS ONE* 10, e0129767.
- Ultsch, A., Lötsch, J. (2017). Machine-learned cluster identification in high-dimensional data. *J Biomed Inform* 66, 95–104.
- Ultsch, A., Sieman, H.P. (1990). Kohonen's self organizing feature maps for exploratory data analysis. *INNOC'90, Int Neural Network Conference (Dordrecht (Netherlands: Kluwer))* 305–308.
- Ultsch, A., Kringel, D., Kalso, E., Mogil, J.S., Lötsch, J. (2016). A data science approach to candidate gene selection of pain regarded as a process of learning and neural plasticity. *Pain* 157, 2747–2757.
- Valdes, A.M., Lories, R.J., van Meurs, J.B., Kerkhof, H., Doherty, S. et al. (2009). Variation at the ANP32A gene is associated with risk of hip osteoarthritis in women. *Arthritis Rheum* 60, 2046–2054.
- Valdes, A.M., Spector, T.D., Tamm, A., Kisand, K., Doherty, S.A. et al. (2010). Genetic variation in the SMAD3 gene is associated with hip and knee osteoarthritis. *Arthritis Rheum* 62, 2347–2352.
- Valdes, A.M., Arden, N.K., Vaughn, F.L., Doherty, S.A., Leaverton, P.E. et al. (2011a). Role of the Nav1.7 R1150W amino acid change in susceptibility to symptomatic knee osteoarthritis and multiple regional pain. *Arthritis Care Res (Hoboken)* 63, 440–444.
- Valdes, A.M., De Wilde, G., Doherty, S.A., Lories, R.J., Vaughn, F.L. et al. (2011b). The Ile585Val TRPV1 variant is involved in risk of painful knee osteoarthritis. *Ann Rheum Dis* 70, 1556–1561.
- Valdes, A.M., Evangelou, E., Kerkhof, H.J., Tamm, A., Doherty, S.A. et al. (2011c). The GDF5 rs143383 polymorphism is associated with osteoarthritis of the knee with genome-wide statistical significance. *Ann Rheum Dis* 70, 873–875.
- Vassalli, P. (1992). The pathophysiology of tumor necrosis factors. *Annu Rev Immunol* 10, 411–452.
- Venn, J. (1880). On the Diagrammatic and Mechanical Representation of Propositions and Reasonings. *Dublin Philos Mag J Sci* 9, 1–18.
- Venter, J.C., Adams, M.D., Myers, E.W., Li, P.W., Mural, R.J. et al. (2001). The sequence of the human genome. *Science* 291, 1304–1351.
- Wells, C.A., Ravasi, T., Faulkner, G.J., Caminci, P., Okazaki, Y. et al. (2003). Genetic control of the innate immune response. *BMC Immunol* 4, 5.
- Witkin, S.S., Gerber, S., Ledger, W.J. (2002). Differential characterization of women with vulvar vestibulitis syndrome. *Am J Obstet Gynecol* 187, 589–594.
- Woo, H.Y., Kim, K.H., Lim, S.W. (2010). Estrogen receptor 1, glutathione S-transferase P1, glutathione S-transferase M1, and glutathione S-transferase T1 genes with dysmenorrhea in Korean female adolescents. *Korean J Lab Med* 30, 76–83.
- Wu, D., Wang, X., Chen, D., Niu, T., Ni, J., Liu, X., Xu, X. (2000). Metabolic gene polymorphisms and risk of dysmenorrhea. *Epidemiology* 11, 648–653.
- Zhang, X., Llamado, L., Pillay, I., Price, P., Will, R. (2002). Interleukin-1 gene polymorphism disease activity and bone mineral metabolism in rheumatoid arthritis. *Chin Med J (Engl)* 115, 46–49.
- Zhang, L., Rao, F., Zhang, K., Khandrika, S., Das, M. et al. (2007). Discovery of common human genetic variants of GTP cyclohydrolase 1 (GCH1) governing nitric oxide, autonomic activity, and cardiovascular risk. *J Clin Invest* 117, 2658–2671.
- Zorina-Lichtenwalter, K., Meloto, C.B., Khoury, S., Diatchenko, L. (2016). Genetic predictors of human chronic pain conditions. *Neuroscience* 338, 36–62.